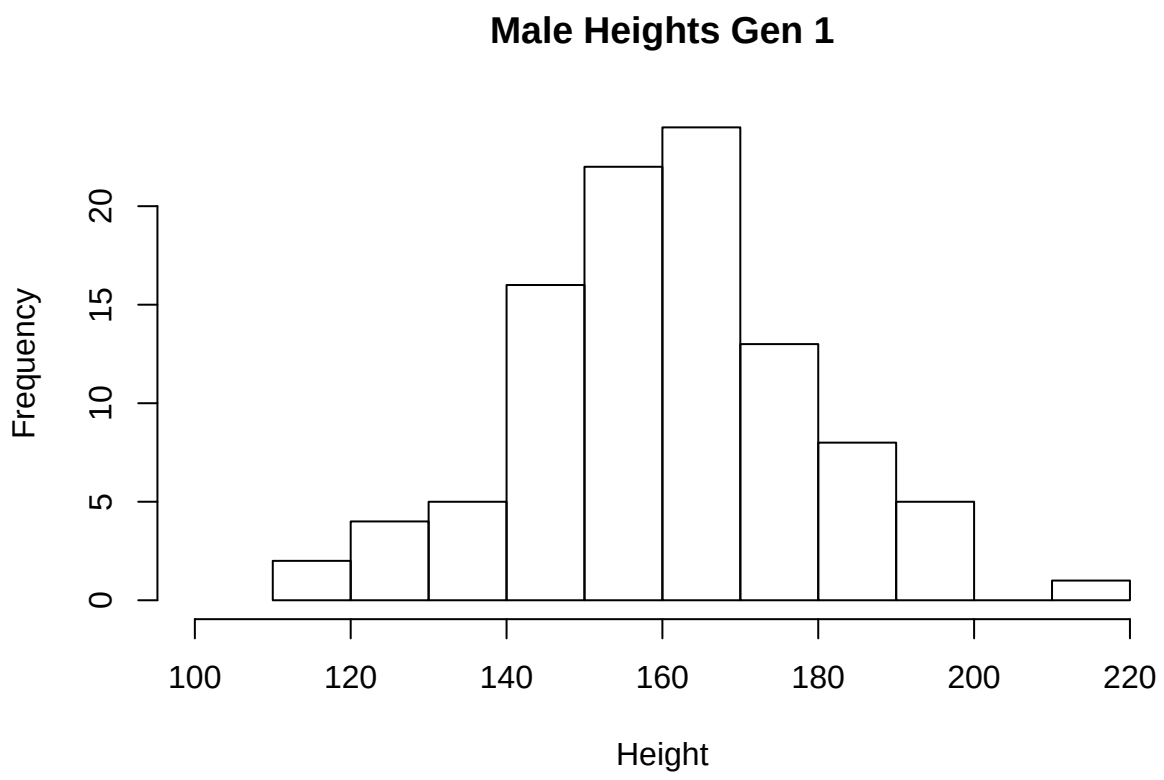
Stat Comp HW 6

Jason Rights

November 23, 2015

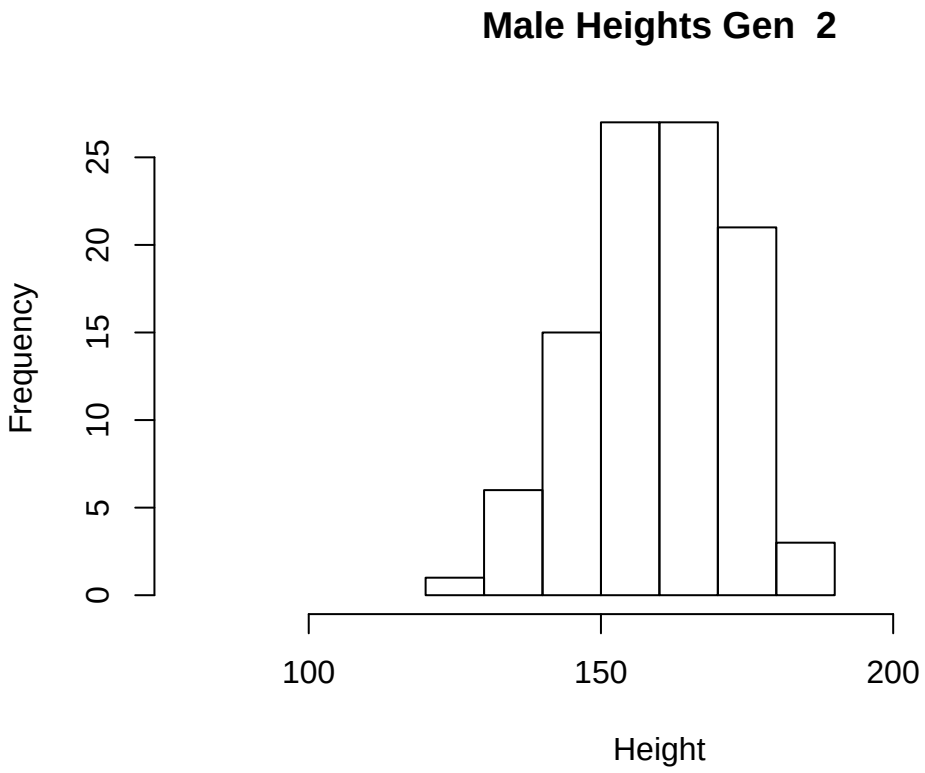
Problem 1:

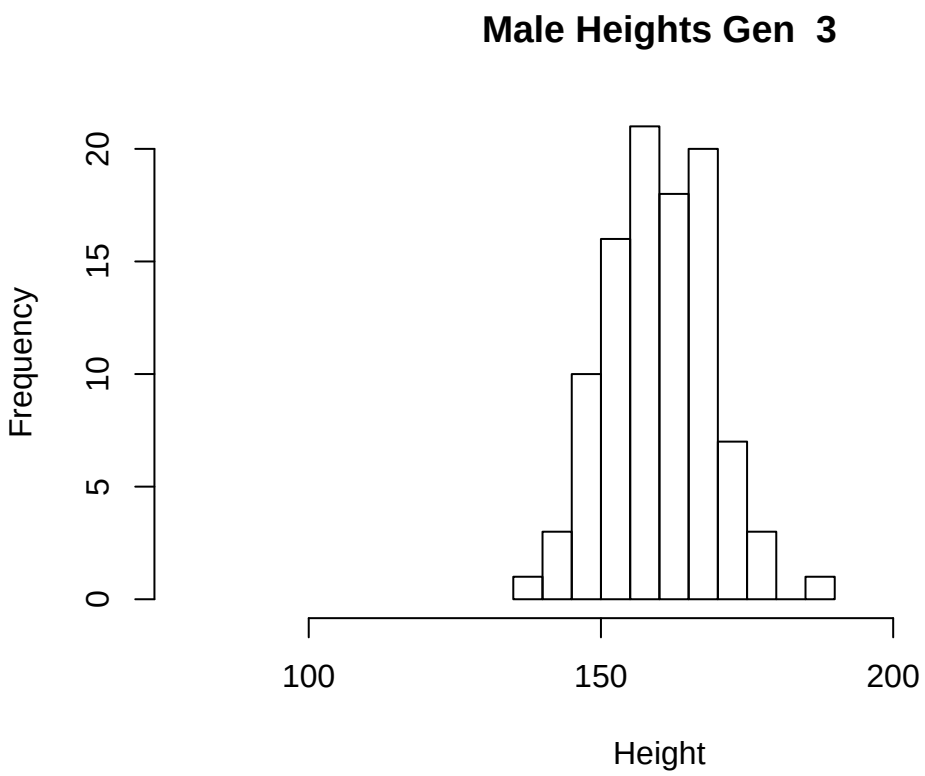
```
pop <- data.frame(m = rnorm(100,160,20), f = rnorm(100, 160, 20))

next_gen <- function(pop) {
    pop$m <- sample(pop$m)
    pop$m <- rowMeans(pop)
    pop$f <- pop$m
    pop
}

hist(pop$m,main='Male Heights Gen 1',xlab='Height',xlim=c(100,220))
```

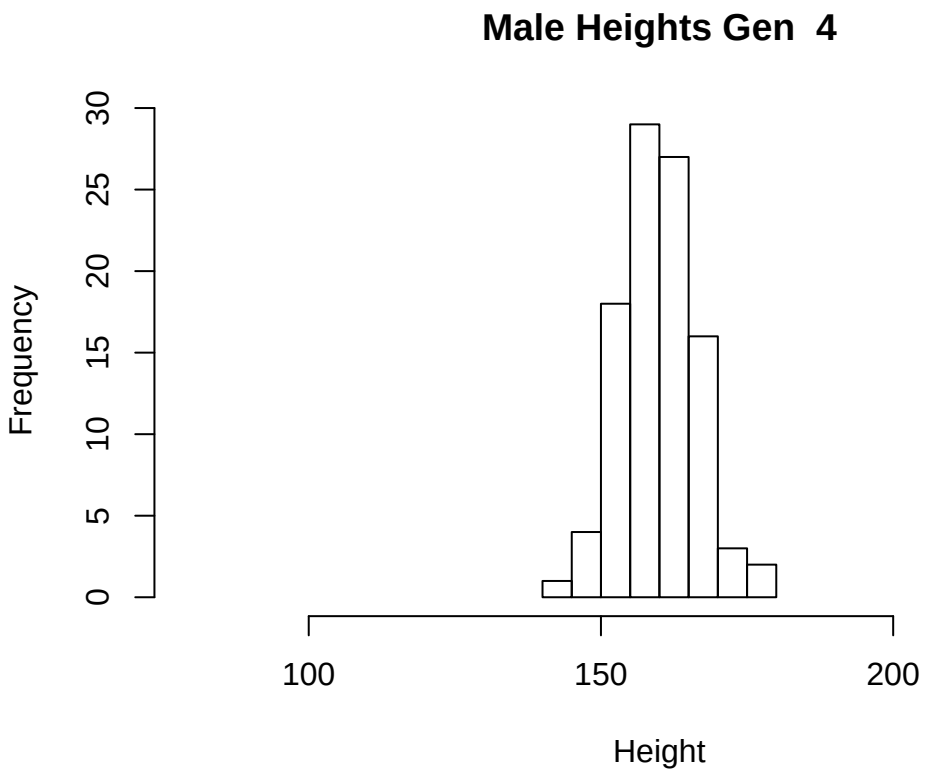


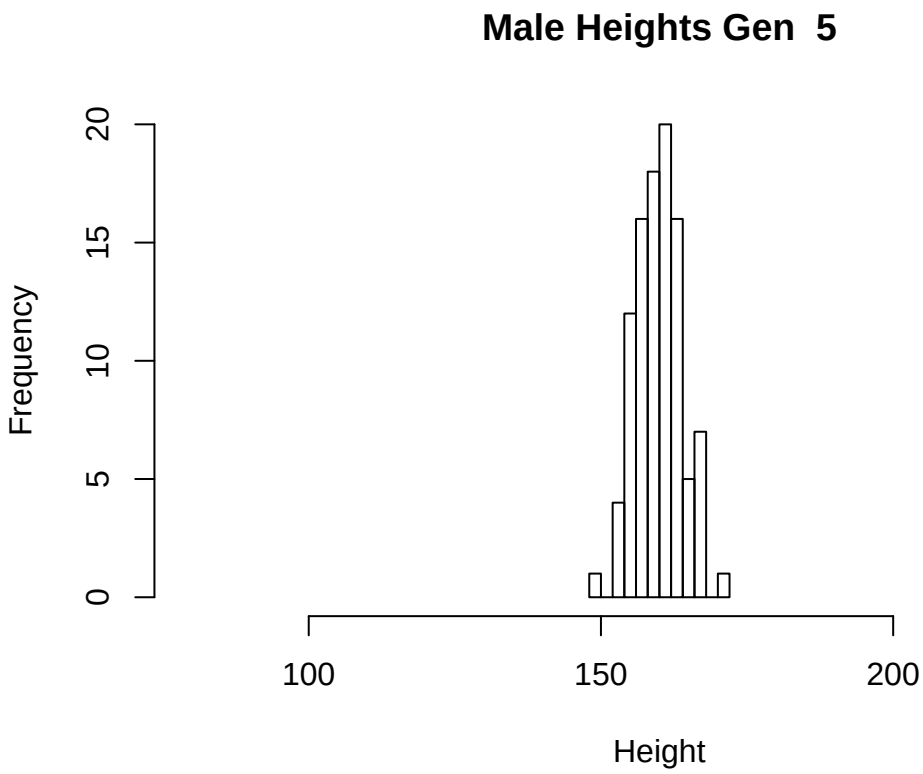

```
next.pop <- as.data.frame(matrix(NA,100,16))
colnames(next.pop) <- rep(c("m","f"),8)
for(i in c(0:7)){
  if (i==0) next.pop[,c(2*i+1):c(2*i+2)] <- next_gen(pop)
  if (i>0) next.pop[,c(2*i+1):c(2*i+2)] <- next_gen(next.pop[,c(2*i-1):c(2*i)])
```

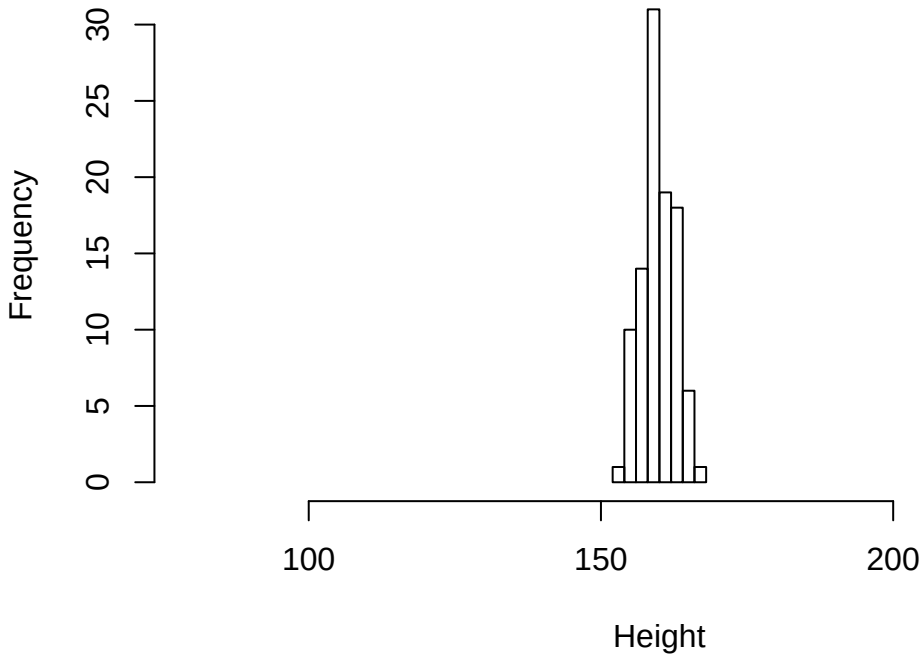
```
  hist(next.pop[,c(2*i+1)],main=paste('Male Heights Gen ',i+2),xlab='Height',xlim=c(80,240))
}
```

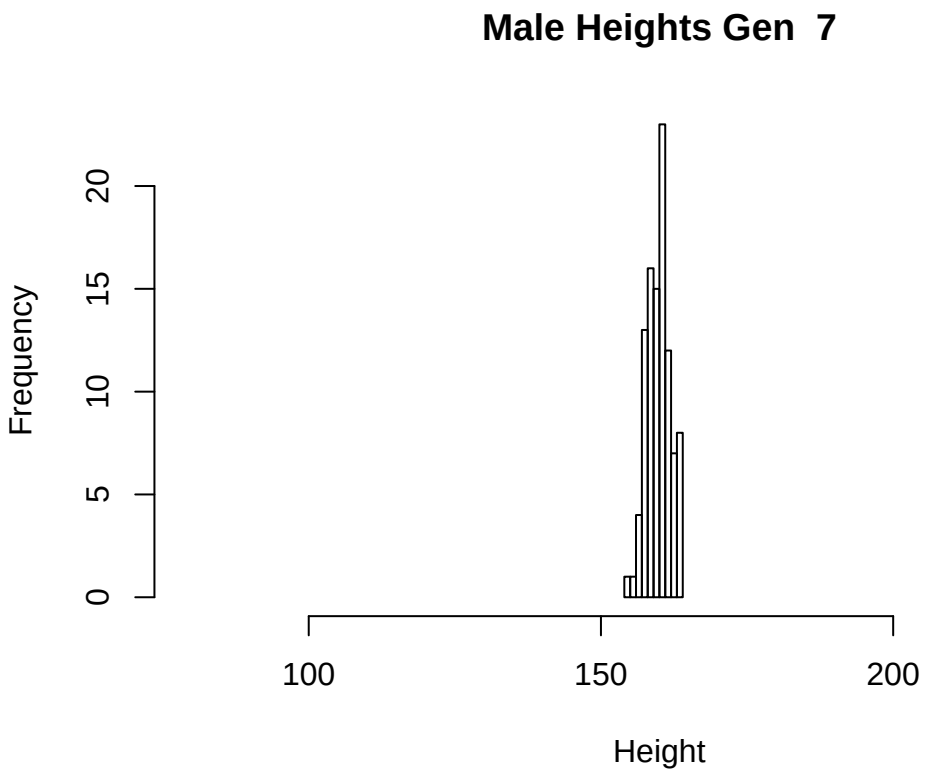


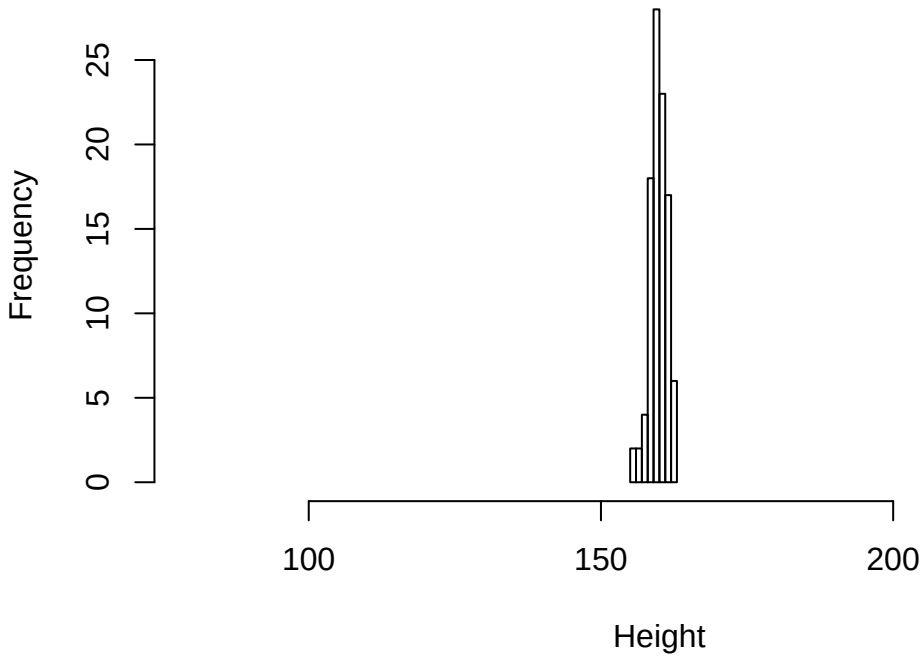

Male Heights Gen 3
Frequency
0
5
10
15
20
100
150
200
Height

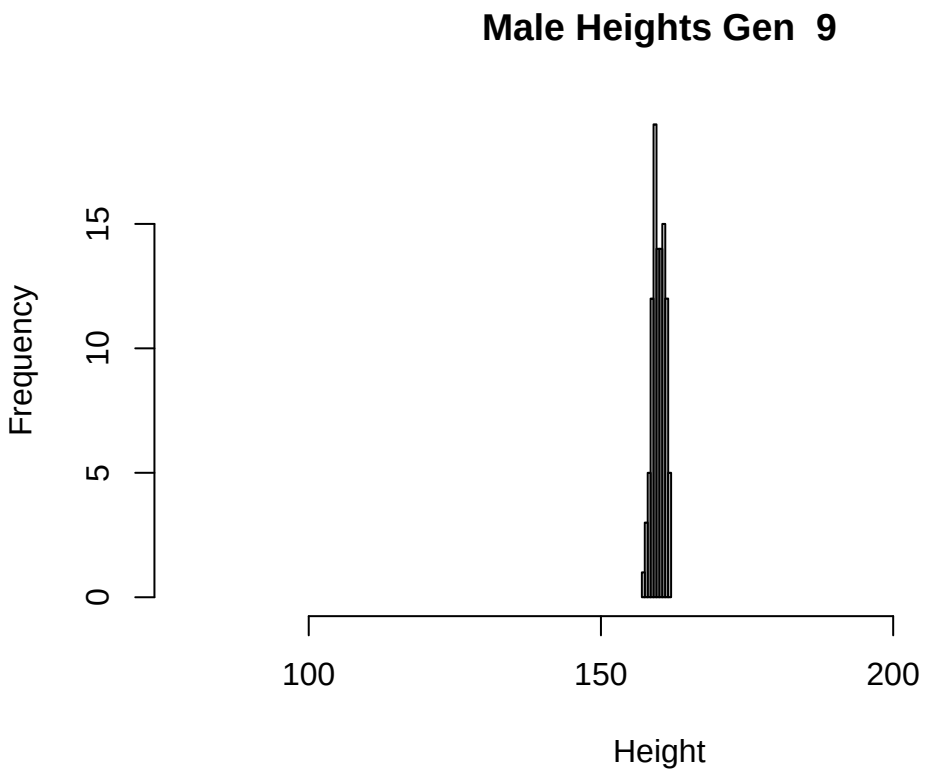

Male Heights Gen 4
Frequency
0
5
10
15
20
25
30
100
150
200
Height


Male Heights Gen 5
Frequency
0
5
10
15
20
100
150
200
Height


Male Heights Gen 6
Frequency
0
5
10
15
20
25
30
100
150
200
Height


Male Heights Gen 7
Frequency
0
5
10
15
20
100
150
200
Height


Male Heights Gen 8
Frequency
0
5
10
15
20
25
100
150
200
Height



I kept the axes the same across all histograms to highlight the effect.

Problem 2:

```
library(ggplot2)
total.pop <- cbind(pop,next.pop)
head(total.pop)
```

```
##          m        f        m        f        m        f        m        f
## 1 190.0696 162.1642 156.1841 156.1841 159.5778 159.5778 154.7926 154.7926
## 2 172.3212 160.3895 159.9736 159.9736 154.9762 154.9762 160.9827 160.9827
## 3 154.4853 177.6219 167.6430 167.6430 158.8772 158.8772 157.3395 157.3395
## 4 162.2287 164.5289 153.0110 153.0110 138.7250 138.7250 153.0829 153.0829
## 5 155.5494 151.7731 151.7752 151.7752 143.8044 143.8044 155.4974 155.4974
## 6 180.3850 196.4489 182.5793 182.5793 177.0693 177.0693 160.7806 160.7806
##          m        f        m        f        m        f        m        f
## 1 157.5076 157.5076 160.9150 160.9150 160.3172 160.3172 158.9843 158.9843
## 2 161.1271 161.1271 159.9338 159.9338 160.4464 160.4464 161.9840 161.9840
## 3 159.1843 159.1843 161.9538 161.9538 162.3386 162.3386 161.7439 161.7439
## 4 154.2901 154.2901 159.5319 159.5319 160.0191 160.0191 159.6030 159.6030
## 5 157.9999 157.9999 159.9907 159.9907 158.7918 158.7918 159.4054 159.4054
## 6 164.5259 164.5259 163.4980 163.4980 161.7284 161.7284 160.0930 160.0930
##          m        f
## 1 159.5691 159.5691
## 2 160.7741 160.7741
## 3 160.2483 160.2483
```

```
## 4 160.6833 160.6833
## 5 160.7252 160.7252
## 6 159.8231 159.8231
```

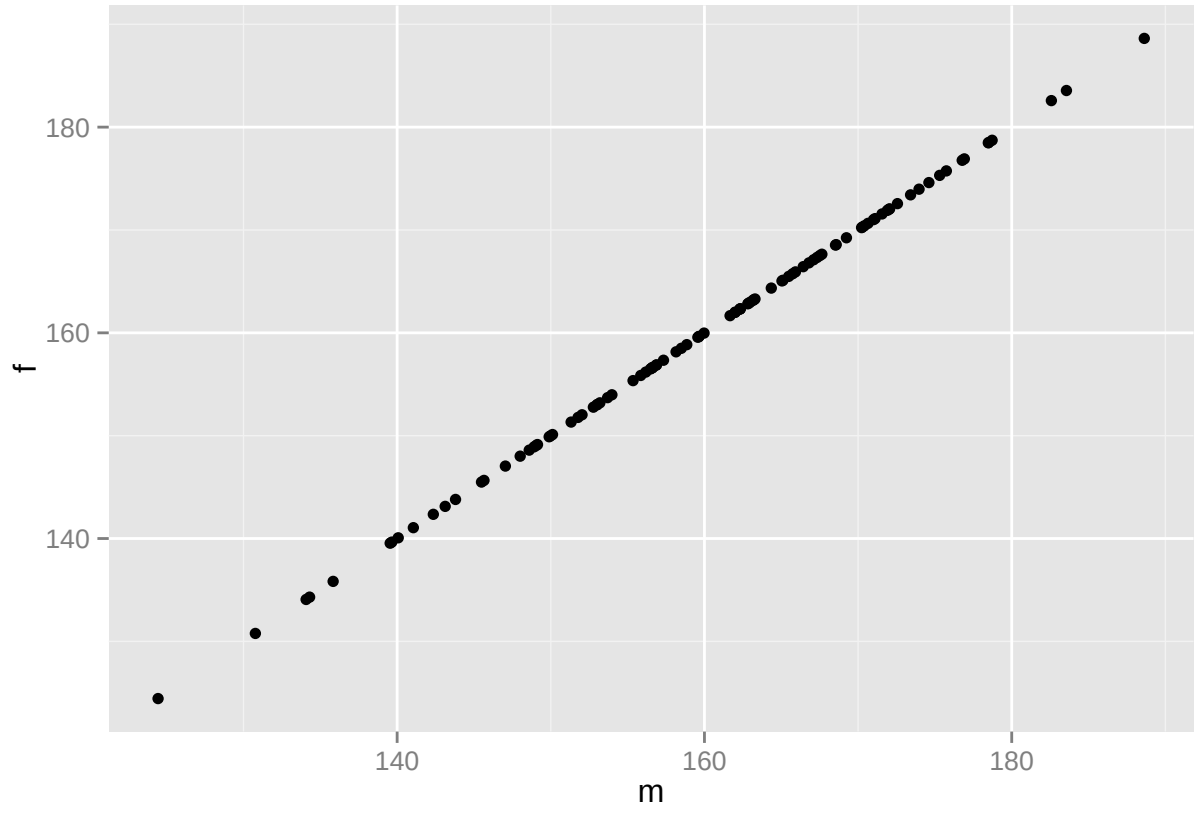
```
dim(total.pop)
```

```
## [1] 100  18
```

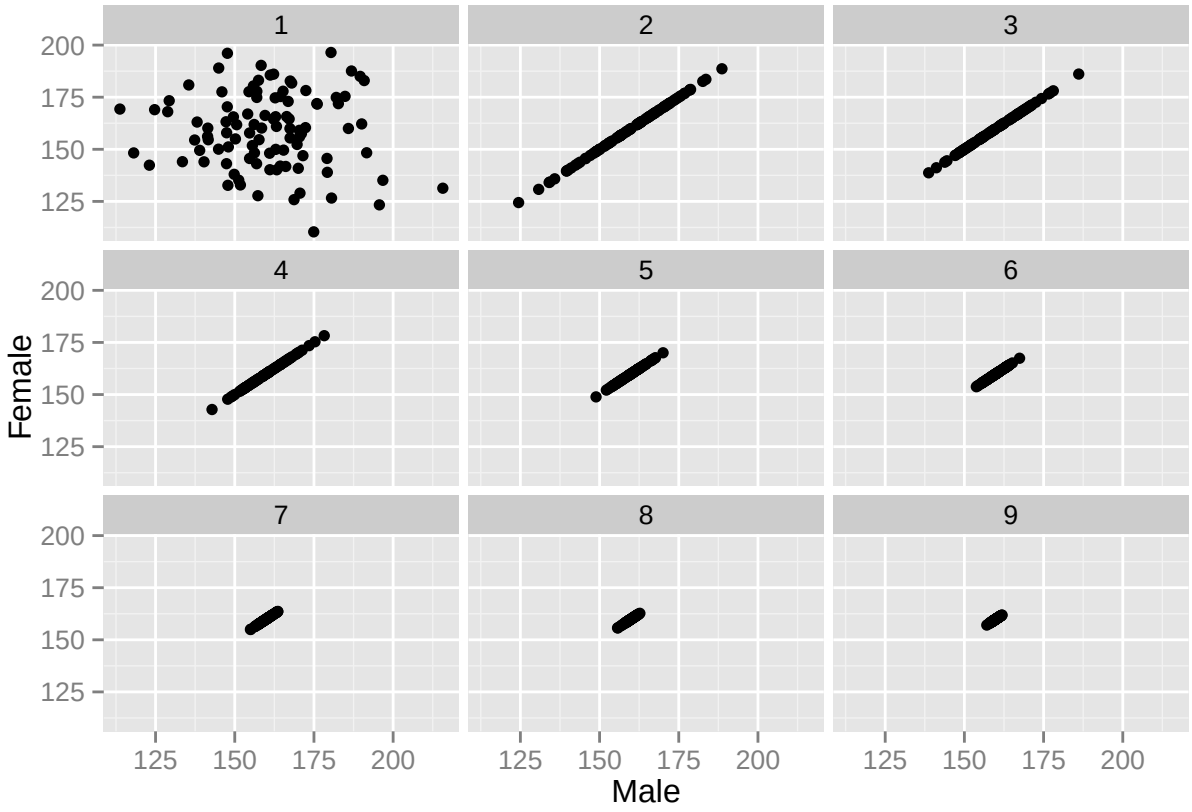
```
total.pop.male <- data.frame(stack(total.pop[,c(1,3,5,7,9,11,13,15,17)]))

total.pop.female <- data.frame(stack(total.pop[,c(2,4,6,8,10,12,14,16,18)]))

total.pop.long <- cbind(total.pop.male[,1],total.pop.female[,1])
total.pop.long <- cbind(total.pop.long,c(rep(1,100),rep(2,100),rep(3,100),rep(4,100),
colnames(total.pop.long) <- c("Male","Female","Gen")
head(total.pop.long)
```

```
##          Male   Female Gen
## [1,] 190.0696 162.1642   1
## [2,] 172.3212 160.3895   1
## [3,] 154.4853 177.6219   1
## [4,] 162.2287 164.5289   1
## [5,] 155.5494 151.7731   1
## [6,] 180.3850 196.4489   1
```

```
qplot(m,f,data=next.pop)
```



```
qplot(Male,Female,data=as.data.frame(total.pop.long),facets=~ Gen)
```



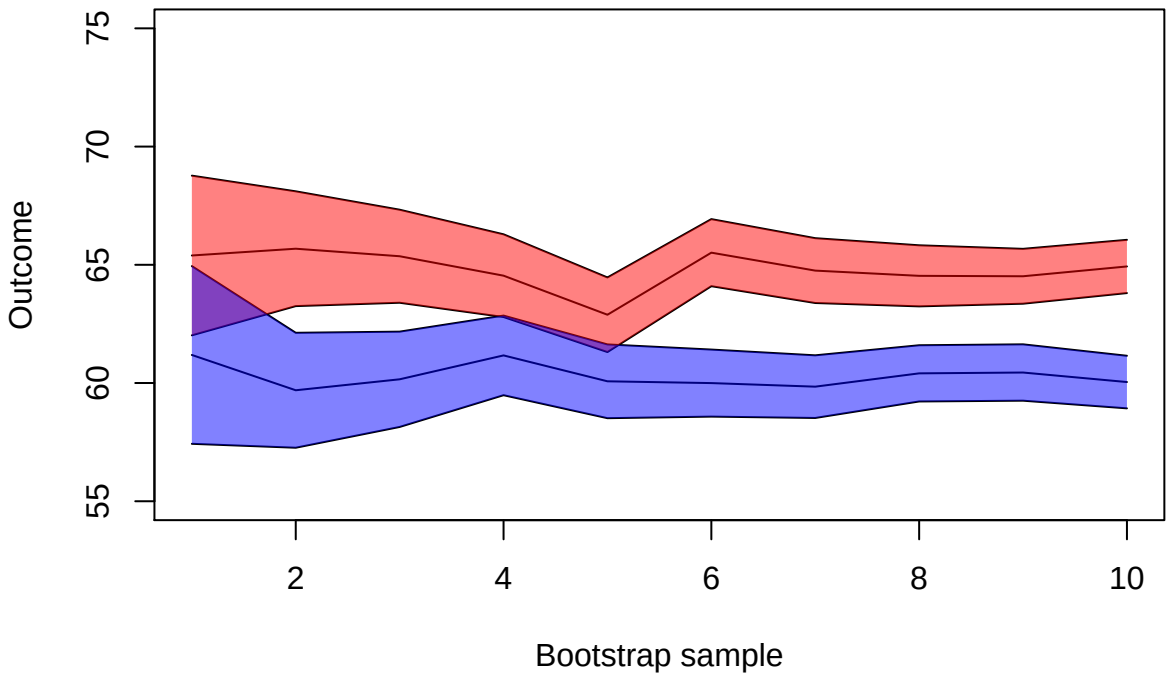
Problem 3:

```
library(boot)

boot.ci.con <- matrix(NA,10,2)
boot.ci.treat <- matrix(NA,10,2)
for(i in seq(10)){
  N <- 250*i
  treatment <- rbinom(N,size=1,prob=.5)
  outcome <- rnorm(N,60,20)
  dat <- as.data.frame(cbind(treatment,outcome))
  for(j in seq(N)){
     dat[j,2] <- ifelse(dat[j,1] == 1, dat[j,2] + 5, dat[j,2])
  }
  dat.con <- subset(dat, treatment==0)
dat.treat <- subset(dat, treatment==1)

  outcome.con_boot <- boot(dat.con$outcome, function(x,i) mean(x[i]), R=2000)
  outcome.treat_boot <- boot(dat.treat$outcome, function(x,i) mean(x[i]), R=2000)

  boot.ci.con[i,] <- boot.ci(outcome.con_boot)$normal[2:3]
  boot.ci.treat[i,] <- boot.ci(outcome.treat_boot)$normal[2:3]
}
```

```
## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals
```

```
## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.con_boot): bootstrap variances needed for
## studentized intervals

## Warning in boot.ci(outcome.treat_boot): bootstrap variances needed for
## studentized intervals
```

```
boot.ci.con <- cbind(boot.ci.con,NA)
boot.ci.treat <- cbind(boot.ci.treat,NA)
for(i in seq(10)){
  boot.ci.con[i,3] <- mean(c(boot.ci.con[i,1],boot.ci.con[i,2]))
  boot.ci.treat[i,3] <- mean(c(boot.ci.treat[i,1],boot.ci.treat[i,2]))
}


plot(boot.ci.con[,1],ylim=c(55,75),col="white",ylab="Outcome",xlab="Bootstrap sample")
lines(boot.ci.con[,1],ylim=c(55,80))
lines(boot.ci.con[,2])
lines(boot.ci.con[,3])
lines(boot.ci.treat[,1],ylim=c(55,80))
lines(boot.ci.treat[,2])
lines(boot.ci.treat[,3])
#legend(100,50,legend=line,lty=c(1,1))

makeTransparent = function(..., alpha=0.5) {
  if(alpha<0 | alpha>1) stop("alpha must be between 0 and 1")
  alpha = floor(255*alpha)
  newColor = col2rgb(col=unlist(list(...)), alpha=FALSE)
  .makeTransparent = function(col, alpha) {
    rgb(red=col[1], green=col[2], blue=col[3], alpha=alpha, maxColorValue=255)
  }
  newColor = apply(newColor, 2, .makeTransparent, alpha=alpha)
  return(newColor)
}

polygon(c(1:10,10:1),
        c(boot.ci.treat[1:10,2],boot.ci.treat[10:1,1]), col=makeTransparent('red',alpha=0.5), border = N
polygon(c(1:10,10:1),
        c(boot.ci.con[1:10,2],boot.ci.con[10:1,1]), col=makeTransparent('blue',alpha=0.5), border = NA)
```




```
#legend("topright", legend = c("Treatment","Control"), col=c("red","blue"), 0lty=c(1:1),lwd=c(3,3),seg.
```

Problem 4:

```
makePatient <- function() {
  vowel <- grep("[aeiou]", letters)
  cons <- grep("[^aeiou]", letters)
  name <- paste(sample(LETTERS[cons], 1), sample(letters[vowel], 1), sample(letters[cons], 1), sep='')
  gender <- factor(sample(0:1, 1), levels=0:1, labels=c('female','male'))
  dob <- as.Date(sample(7500, 1), origin="1970-01-01")
  n <- sample(6, 1)
  doa <- as.Date(sample(1500, n), origin="2010-01-01")
  pulse <- round(rnorm(n, 80, 10))
  temp <- round(rnorm(n, 98.4, 0.3), 2)
  fluid <- round(runif(n), 2)
  x <- list(name, gender, dob, doa, pulse, temp, fluid)
  names(x) <- list("name", "gender", "date_of_birth", "date_of_admission", "pulse", "temperature", "flu
  class(x) <- "medicalRecord"
  x
}
set.seed(8)
med.record <- makePatient()
med.record
```

```
## $name
```

```
## [1] "Mev"
##
## $gender
## [1] male
## Levels: female male
##
## $date_of_birth
## [1] "1976-08-09"
##
## $date_of_admission
## [1] "2011-03-14" "2013-10-30" "2013-02-27" "2012-08-23" "2011-11-16"
##
## $pulse
## [1] 67 81 95 74 81
##
## $temperature
## [1] 98.33 98.16 99.00 98.49 98.67
##
## $fluid_intake
## [1] 0.62 0.93 0.18 0.39 0.34
##
## attr(,"class")
## [1] "medicalRecord"
```

```
class(med.record)
```

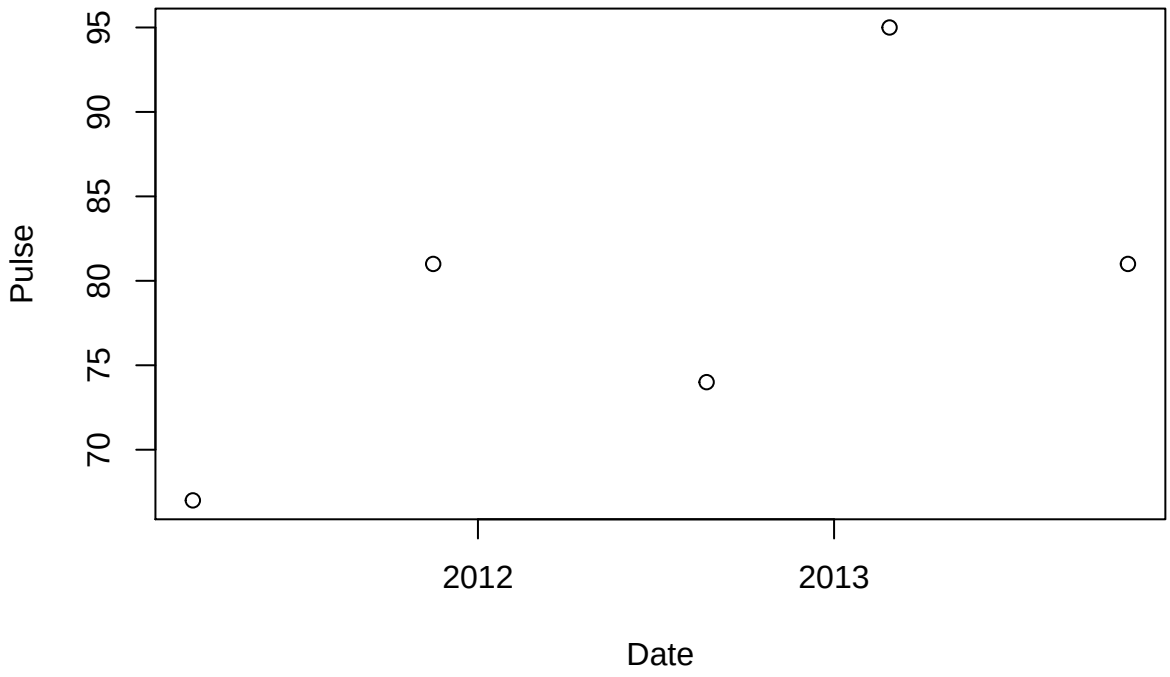
```
## [1] "medicalRecord"
```

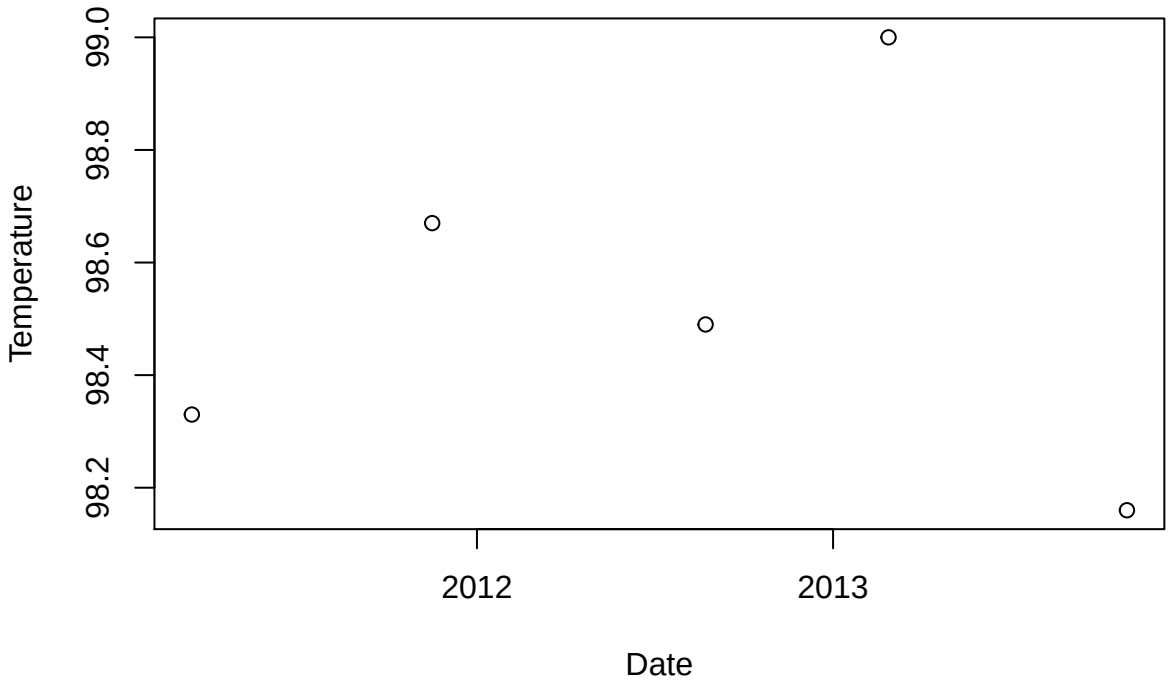
```
mean.medicalRecord <- function(med){
  x <- list(mean(med$pulse),mean(med$temperature),mean(med$fluid_intake))
  names(x) <- c("mean_pulse","mean_temperature","mean_fluid_intake")
  return(x)
}

mean(med.record)
```

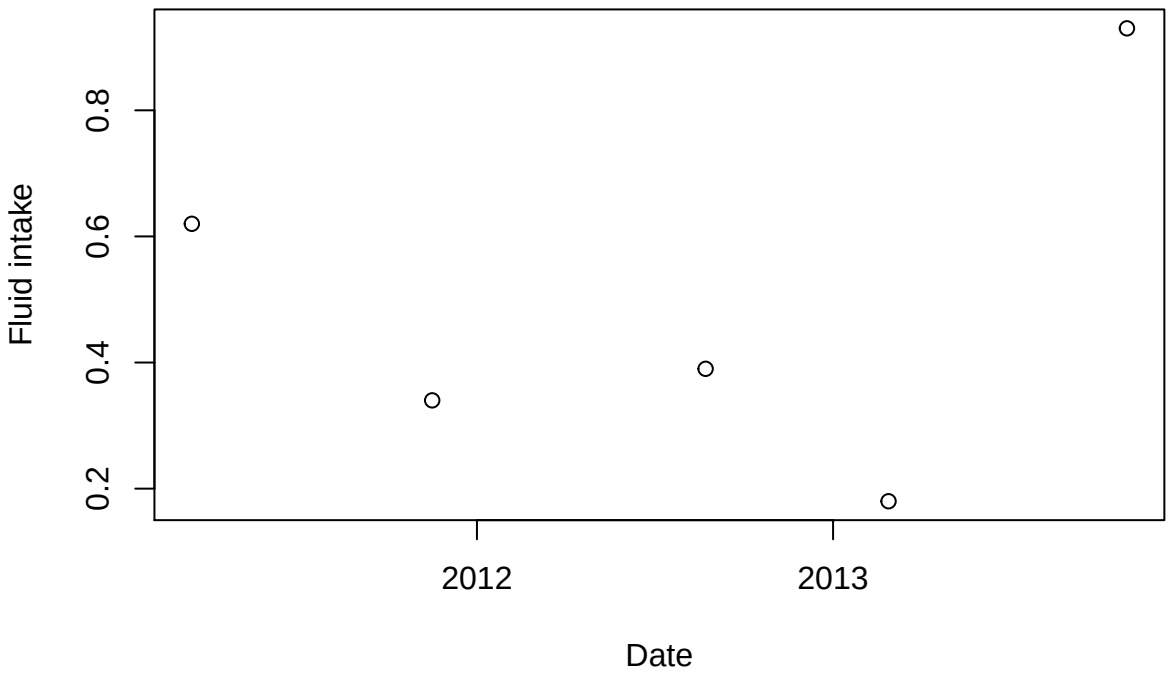
```
## $mean_pulse
## [1] 79.6
##
## $mean_temperature
## [1] 98.53
##
## $mean_fluid_intake
## [1] 0.492
```

```
print.medicalRecord <- function(med){
  x <- data.frame(med$date_of_admission,med$pulse,med$temperature,med$fluid_intake)
  x <- x[order(med$date_of_admission),]
  colnames(x) <- c("Date of admission","Pulse","Temperature","Fluid intake")
  rownames(x) <- NULL
  plot(x[,1],x[,2],xlab="Date",ylab="Pulse")
  plot(x[,1],x[,3],xlab="Date",ylab="Temperature")
```

```
  plot(x[,1],x[,4],xlab="Date",ylab="Fluid intake")
  return(x)
}

print(med.record)
```




Temperature
98.2
98.4
98.6
98.8
99.0
2012
2013
Date




```
##   Date of admission Pulse Temperature Fluid intake
## 1        2011-03-14    67       98.33         0.62
## 2        2011-11-16    81       98.67         0.34
## 3        2012-08-23    74       98.49         0.39
## 4        2013-02-27    95       99.00         0.18
## 5        2013-10-30    81       98.16         0.93
```

```
##part 3

makePatientCohort <- function(n.patients) {
  cohort <- makePatient()
  for(i in c(2:n.patients)){
    cohort <- c(cohort,makePatient())
  }
  class(cohort) <- c("cohort")
   return(cohort)
}


x <-makePatientCohort(10)
class(x)
```

```
## [1] "cohort"
```

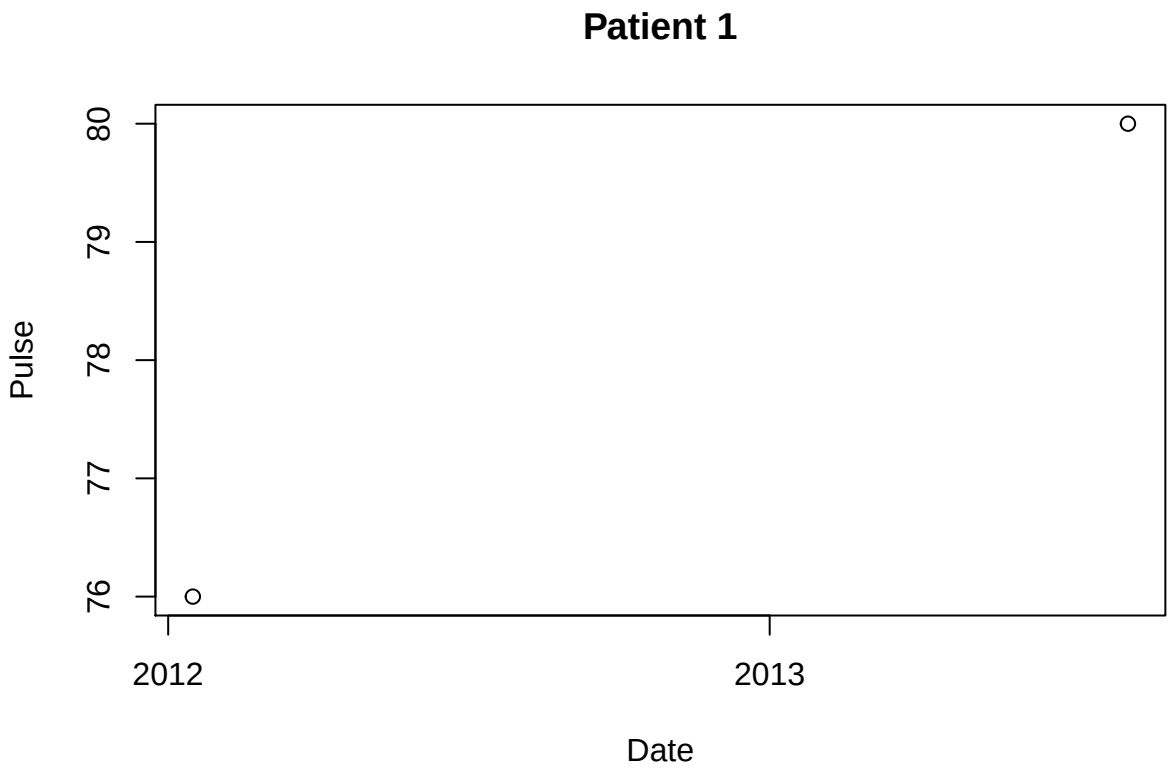
```
mean.cohort <- function(cohort){
  for(i in seq(length(cohort)/7)){
  x <- list(mean(cohort[c(7*(i-1)+1):c(7*i)]$pulse),mean(cohort[c(7*(i-1)+1):c(7*i)]$temperature),mean(
  names(x) <- c(paste("mean_pulse_patient_",i),paste("mean_temperature_patient_",i),paste("mean_fluid_in
 print(x)
   }

}


print.cohort <- function(cohort){
  for(i in seq(length(cohort)/7)){
  x <- data.frame(cohort[c(7*(i-1)+1):c(7*i)]$date_of_admission,cohort[c(7*(i-1)+1):c(7*i)]$pulse,cohort
  x <- x[order(cohort[c(7*(i-1)+1):c(7*i)]$date_of_admission),]
  colnames(x) <- c(paste("Date of admission, patient",i),"Pulse","Temperature","Fluid intake")
  rownames(x) <- NULL
  plot(x[,1],x[,2],xlab="Date",ylab="Pulse",main=paste("Patient",i))
  plot(x[,1],x[,3],xlab="Date",ylab="Temperature",main=paste("Patient",i))
  plot(x[,1],x[,4],xlab="Date",ylab="Fluid intake",main=paste("Patient",i))
  print(x)
  }
}

mean(x)
```

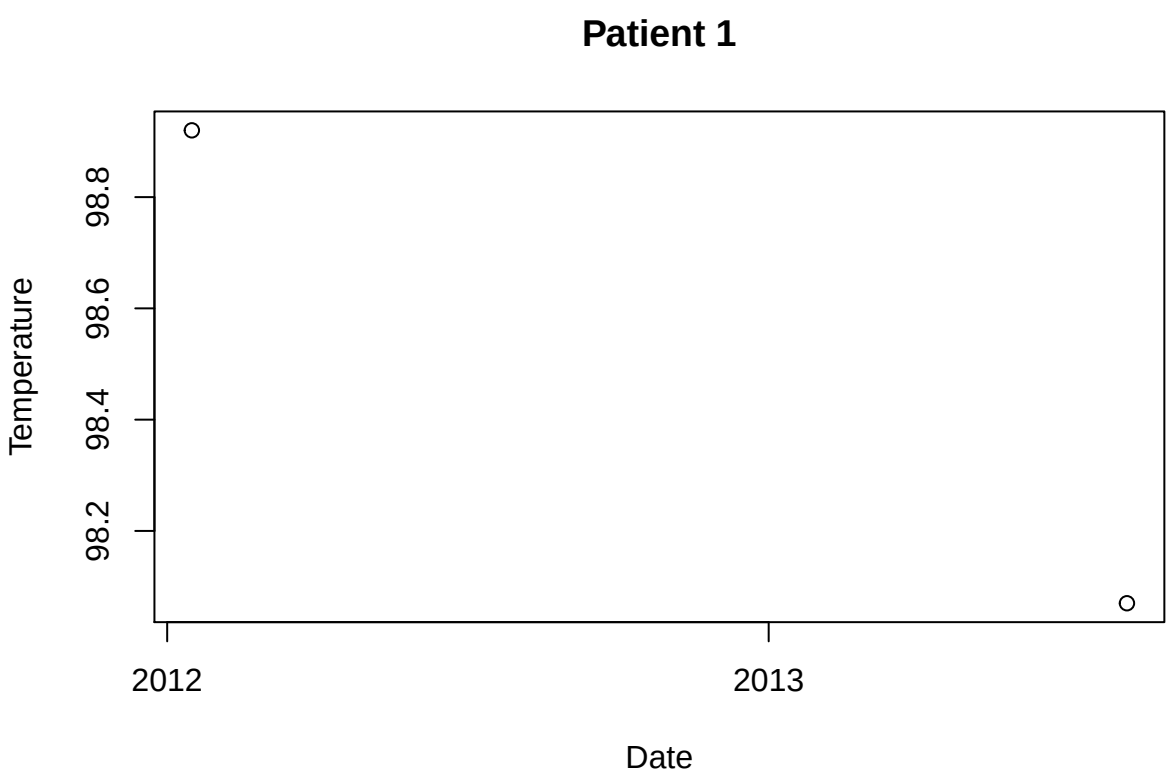
```
## $`mean_pulse_patient_ 1`
## [1] 78
##
## $`mean_temperature_patient_ 1`
## [1] 98.495
##
## $`mean_fluid_intake_patient_ 1`
## [1] 0.245
##
## $`mean_pulse_patient_ 2`
## [1] 81.5
##
## $`mean_temperature_patient_ 2`
## [1] 98.44
##
## $`mean_fluid_intake_patient_ 2`
## [1] 0.4033333
##
## $`mean_pulse_patient_ 3`
## [1] 78
##
## $`mean_temperature_patient_ 3`
## [1] 98.6
##
## $`mean_fluid_intake_patient_ 3`
## [1] 0.65
##
## $`mean_pulse_patient_ 4`
```

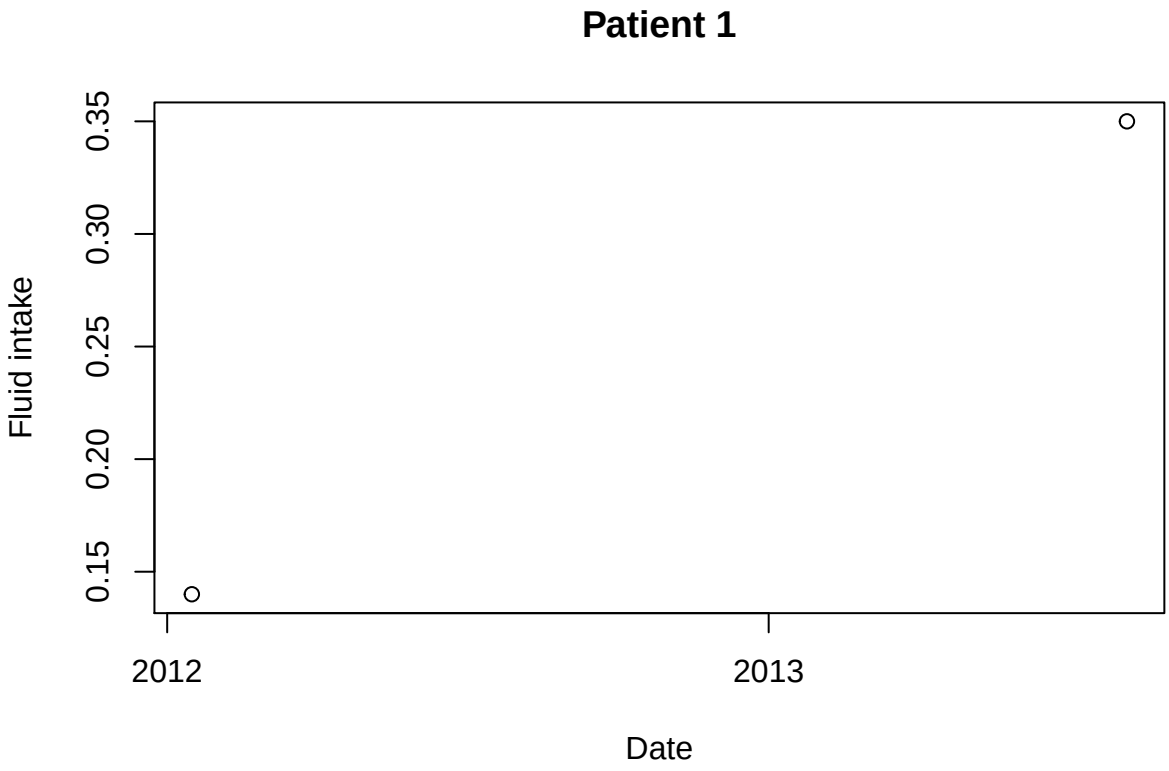
```
## [1] 88.33333
##
## $`mean_temperature_patient_ 4`
## [1] 98.05
##
## $`mean_fluid_intake_patient_ 4`
## [1] 0.5866667
##
## $`mean_pulse_patient_ 5`
## [1] 83.5
##
## $`mean_temperature_patient_ 5`
## [1] 98.45
##
## $`mean_fluid_intake_patient_ 5`
## [1] 0.4525
##
## $`mean_pulse_patient_ 6`
## [1] 83
##
## $`mean_temperature_patient_ 6`
## [1] 98.01
##
## $`mean_fluid_intake_patient_ 6`
## [1] 0.97
##
## $`mean_pulse_patient_ 7`
## [1] 77.5
##
## $`mean_temperature_patient_ 7`
## [1] 98.14833
##
## $`mean_fluid_intake_patient_ 7`
## [1] 0.3366667
##
## $`mean_pulse_patient_ 8`
## [1] 77
##
## $`mean_temperature_patient_ 8`
## [1] 98.83
##
## $`mean_fluid_intake_patient_ 8`
## [1] 0.445
##
## $`mean_pulse_patient_ 9`
## [1] 79.33333
##
## $`mean_temperature_patient_ 9`
## [1] 98.3
##
## $`mean_fluid_intake_patient_ 9`
## [1] 0.6583333
##
## $`mean_pulse_patient_ 10`
```

```
## [1] 81
##
## $`mean_temperature_patient_ 10`
## [1] 98.21
##
## $`mean_fluid_intake_patient_ 10`
## [1] 0.64
```

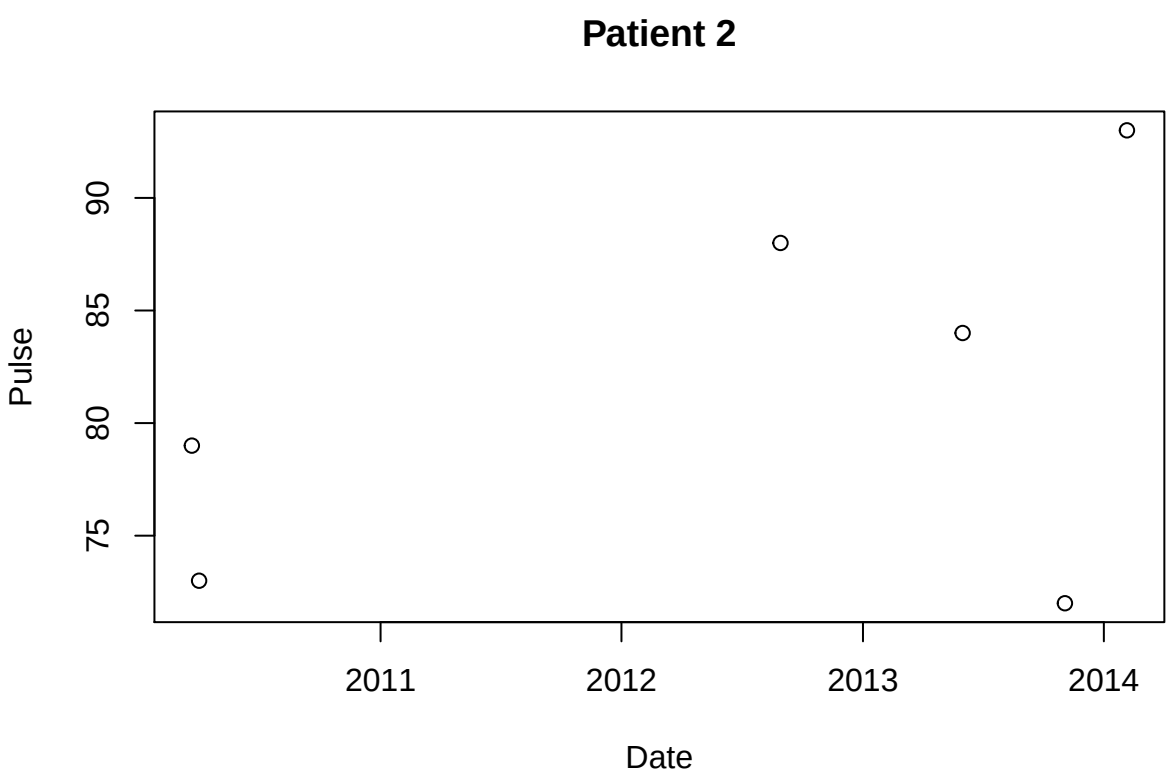
```
print(x)
```

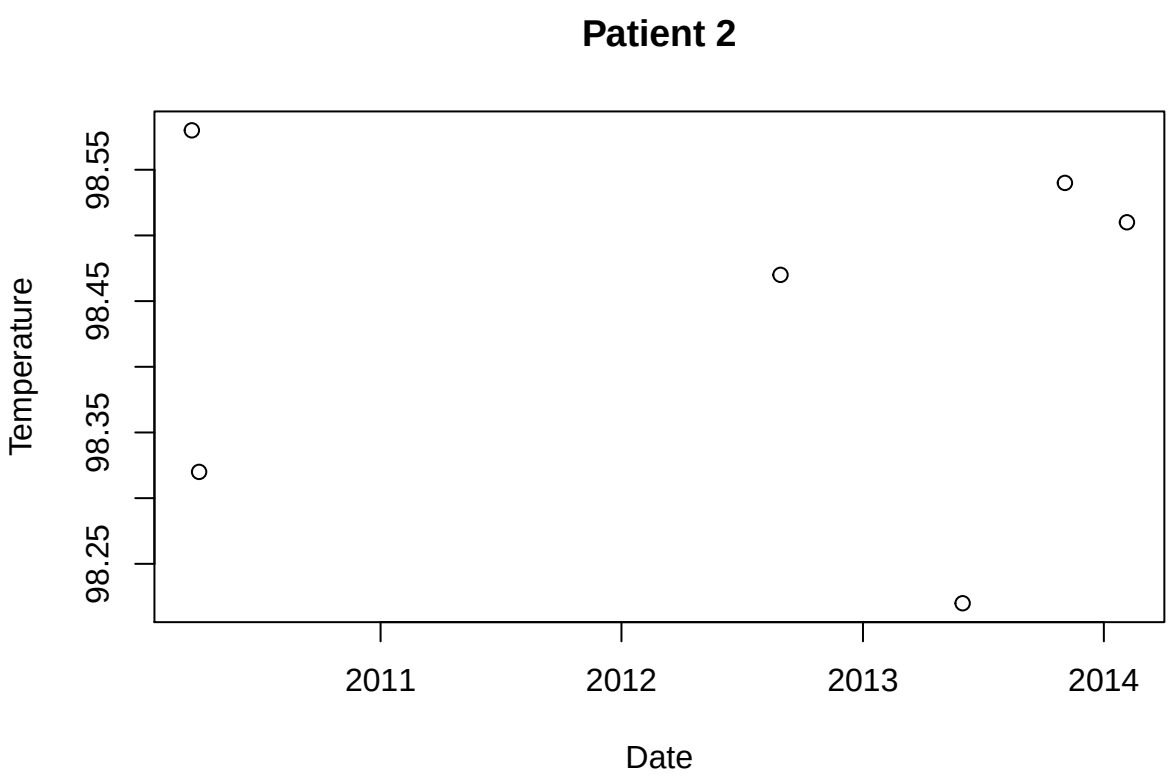


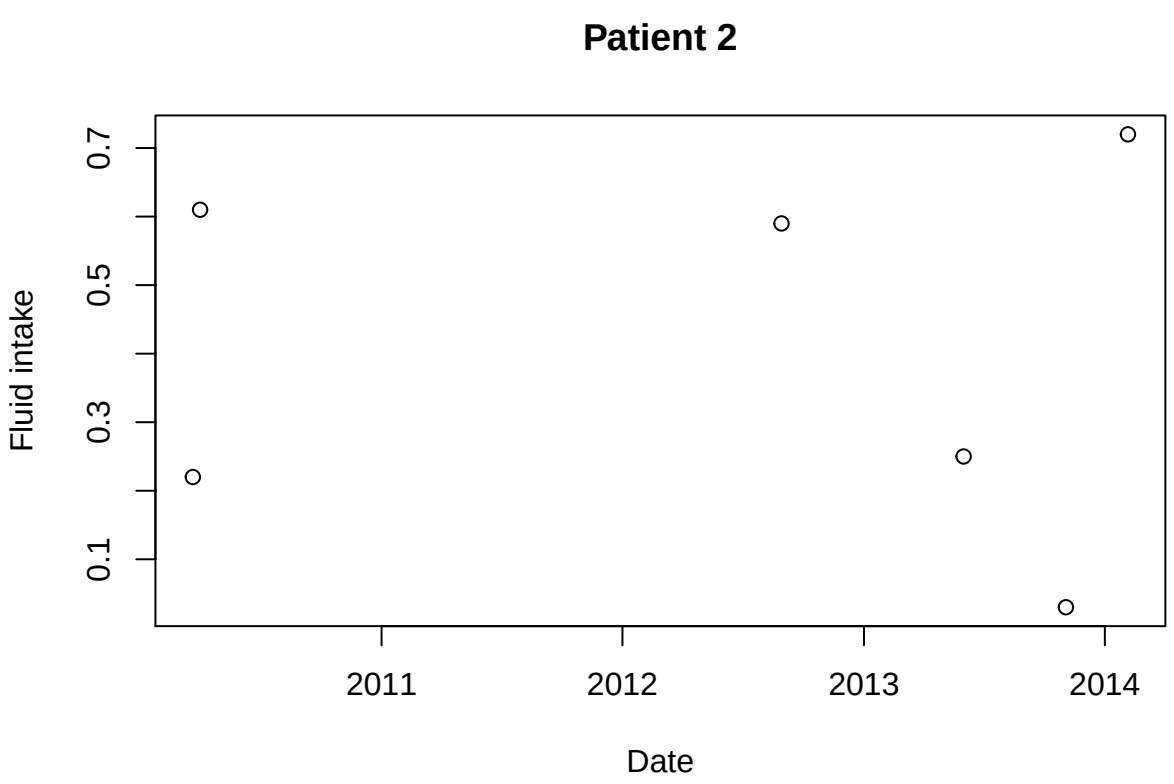

Patient 1
Temperature
98.2
98.4
98.6
98.8
2012
2013
Date



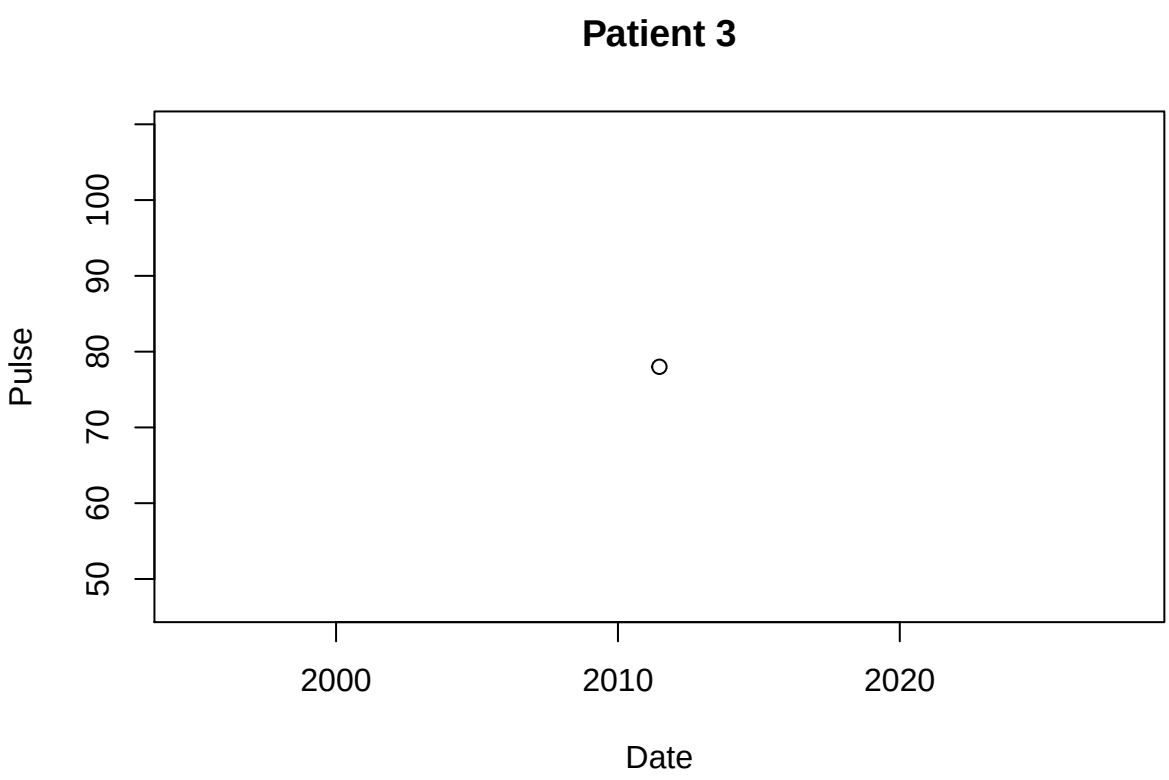

```
##   Date of admission, patient 1 Pulse Temperature Fluid intake
## 1                   2012-01-16    76       98.92         0.14
## 2                   2013-08-07    80       98.07         0.35
```

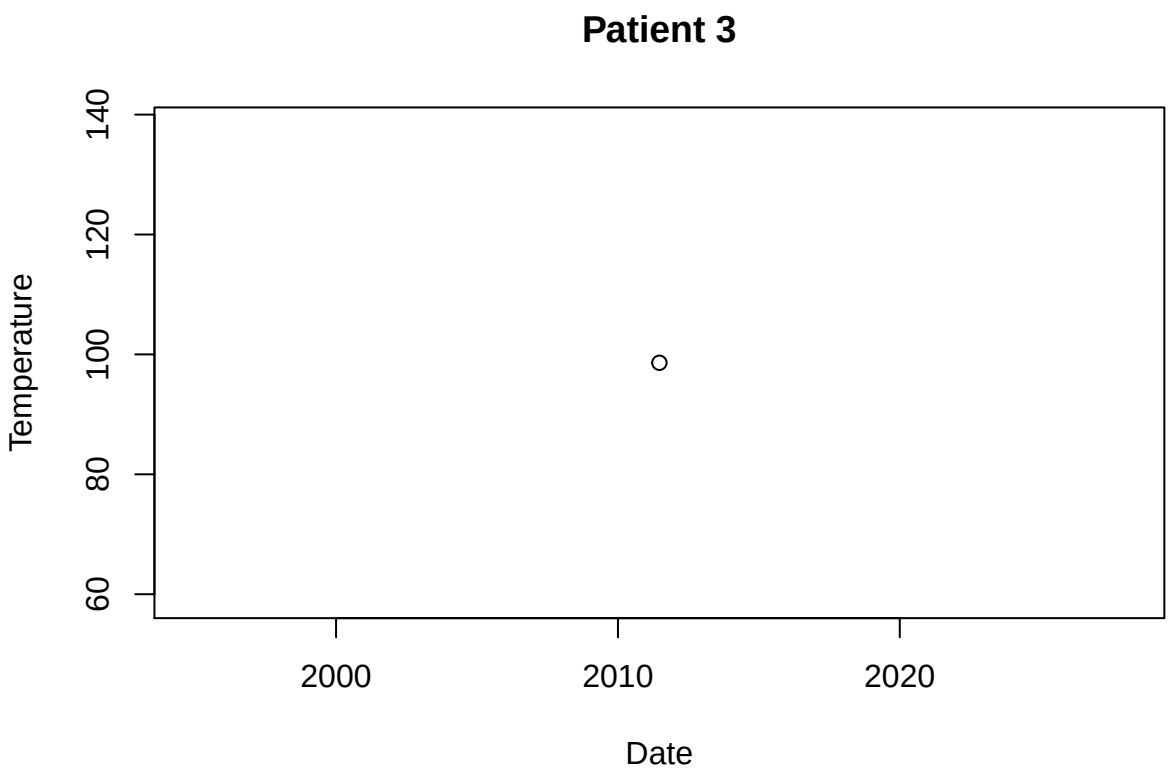

Patient 2
Pulse
75
80
85
90
2011
2012
2013
2014
Date

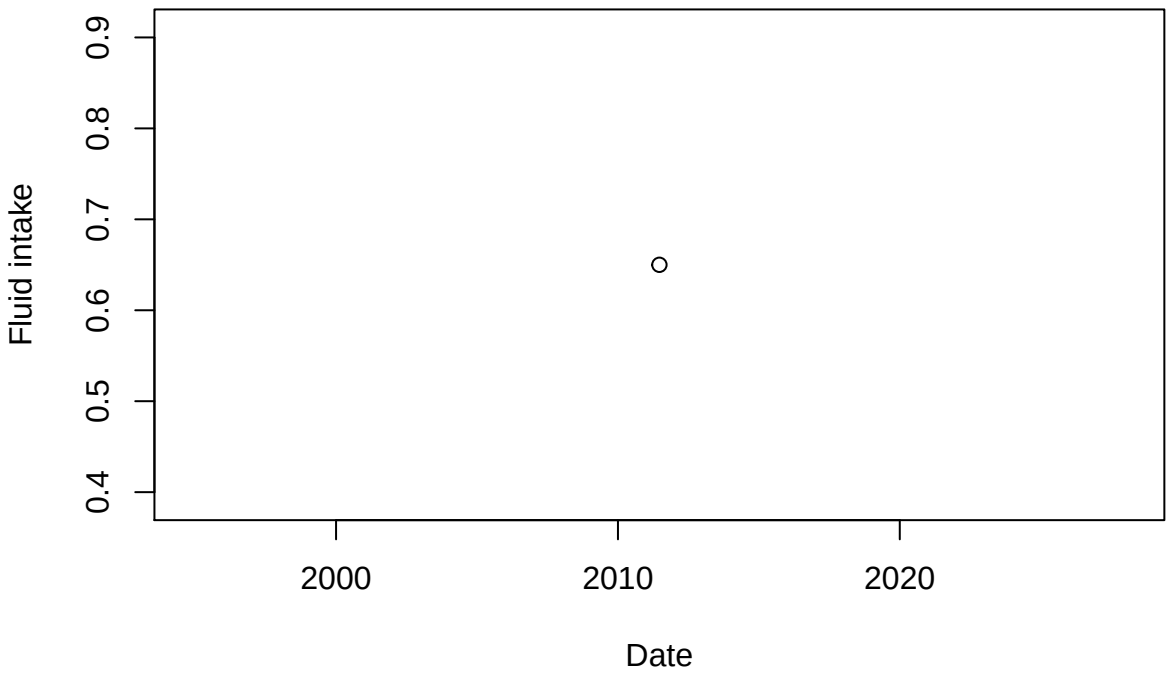

Patient 2
Temperature
98.25
98.35
98.45
98.55
2011
2012
2013
2014
Date



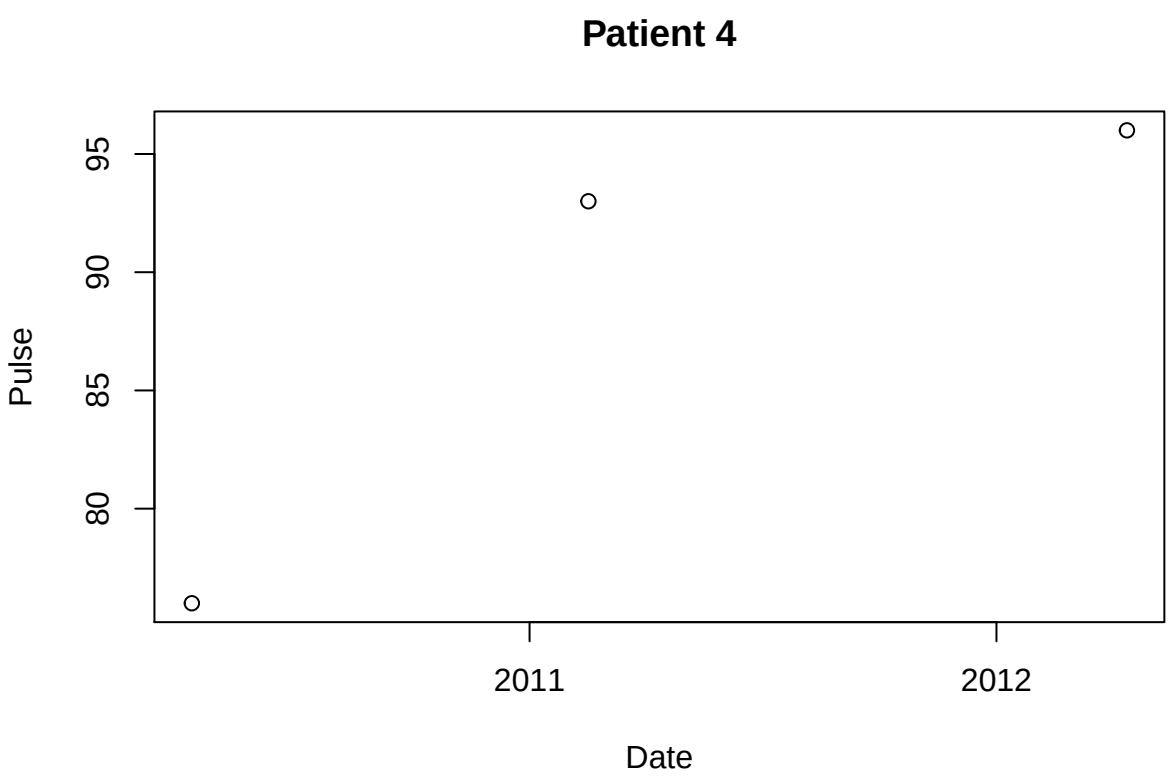

```
##   Date of admission, patient 2 Pulse Temperature Fluid intake
## 1                   2010-03-21    79       98.58         0.22
## 2                   2010-04-01    73       98.32         0.61
## 3                   2012-08-29    88       98.47         0.59
## 4                   2013-06-01    84       98.22         0.25
## 5                   2013-11-03    72       98.54         0.03
## 6                   2014-02-05    93       98.51         0.72
```

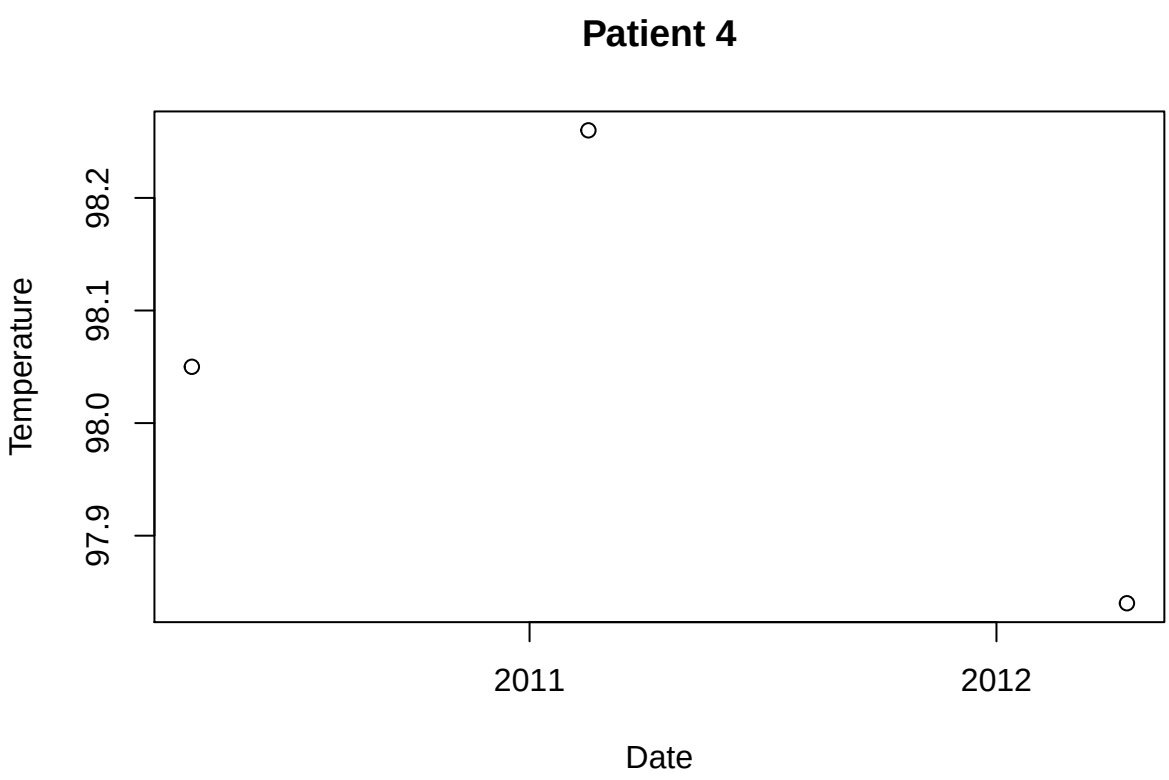

Patient 3
Pulse
50
60
70
80
90
100
2000
2010
2020
Date

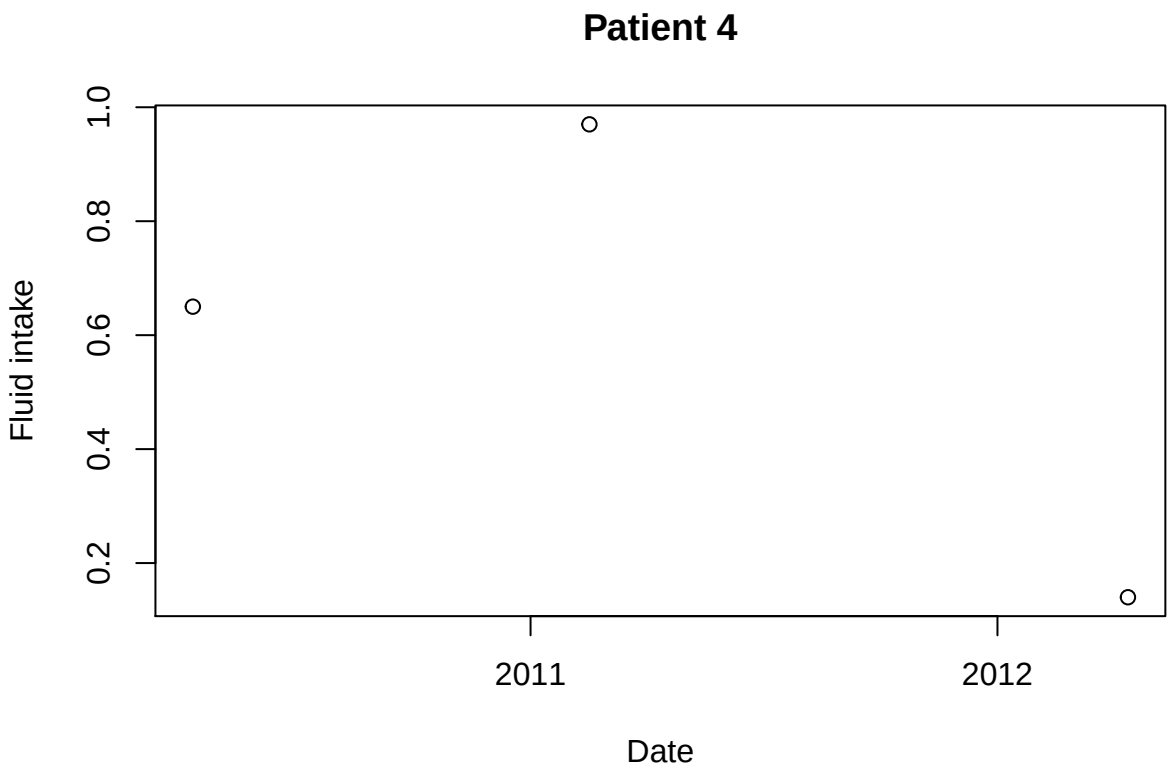

Patient 3
Temperature
140
120
100
80
60
2000
2010
2020
Date



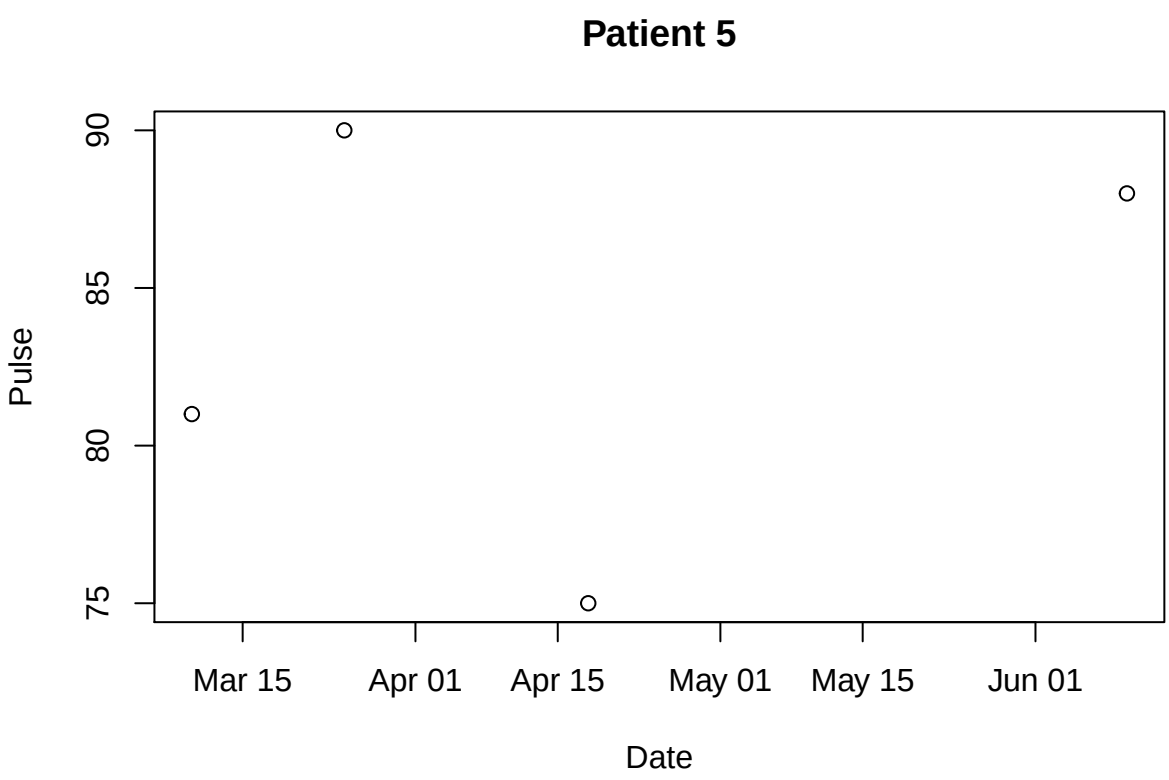

```
##   Date of admission, patient 3 Pulse Temperature Fluid intake
## 1                   2011-06-22    78        98.6         0.65
```

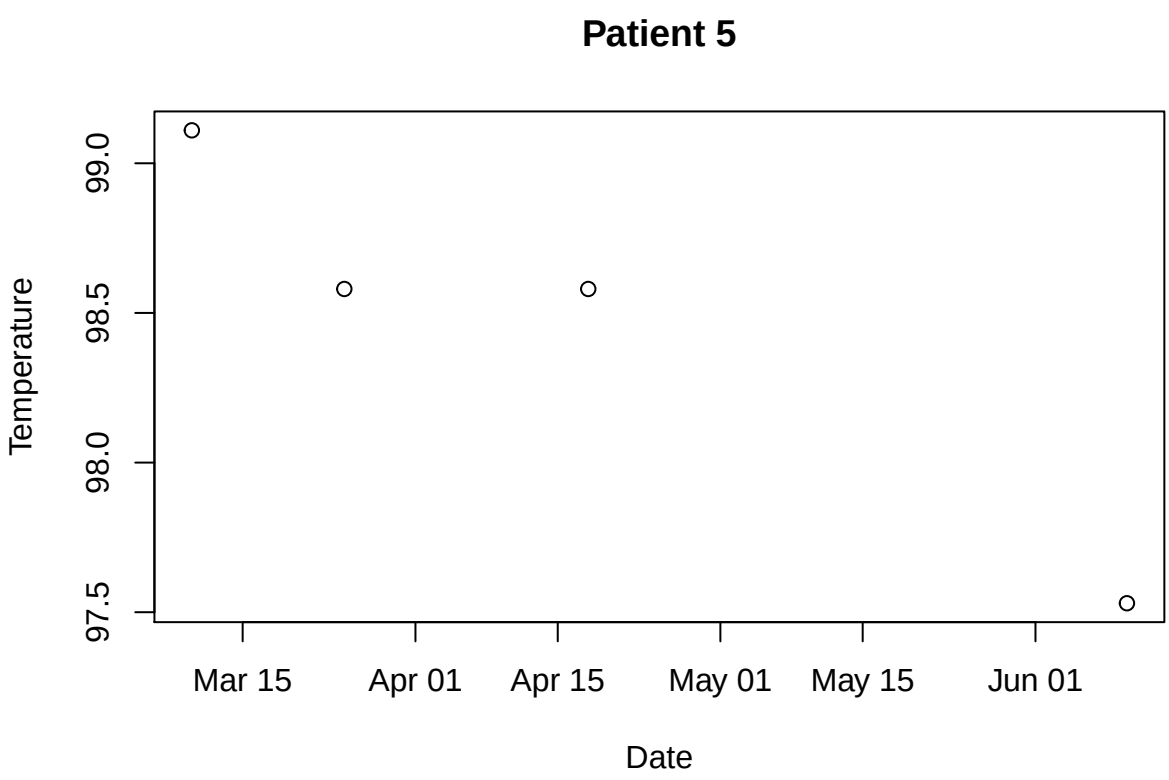

Patient 4
Pulse
80
85
90
95
2011
2012
Date

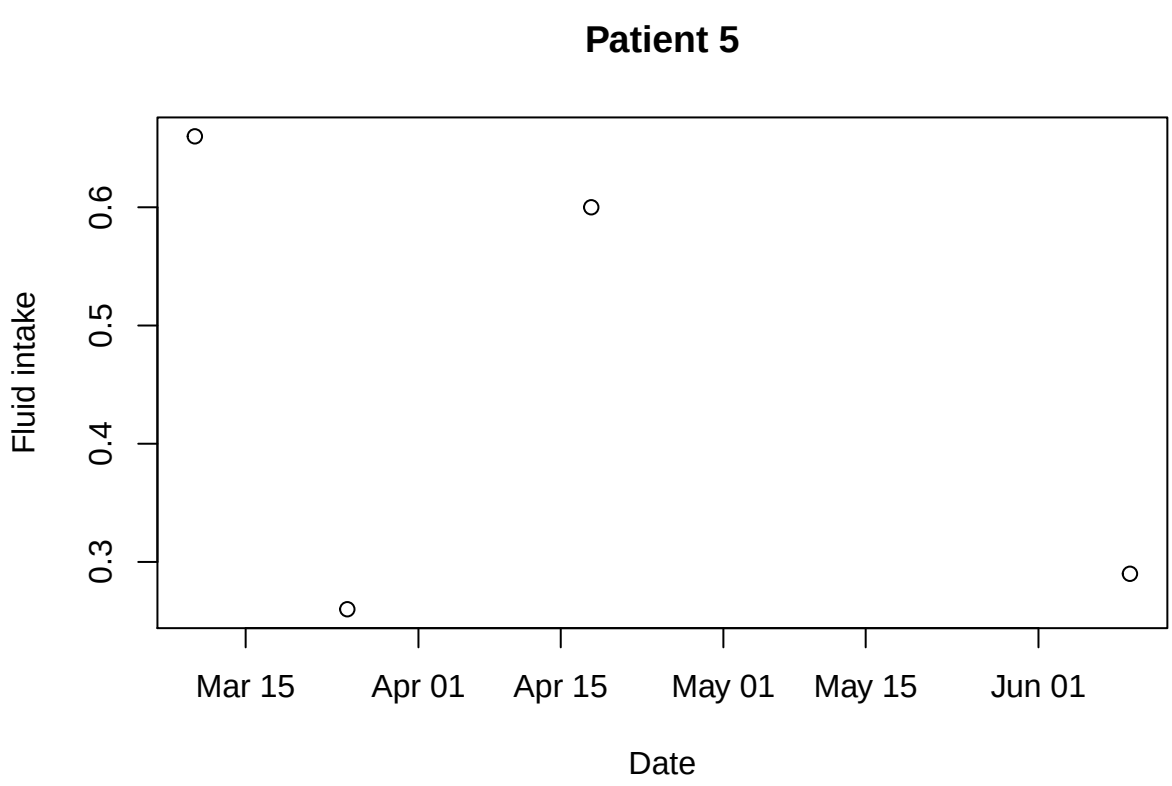

Patient 4
Temperature
97.9
98.0
98.1
98.2
2011
2012
Date



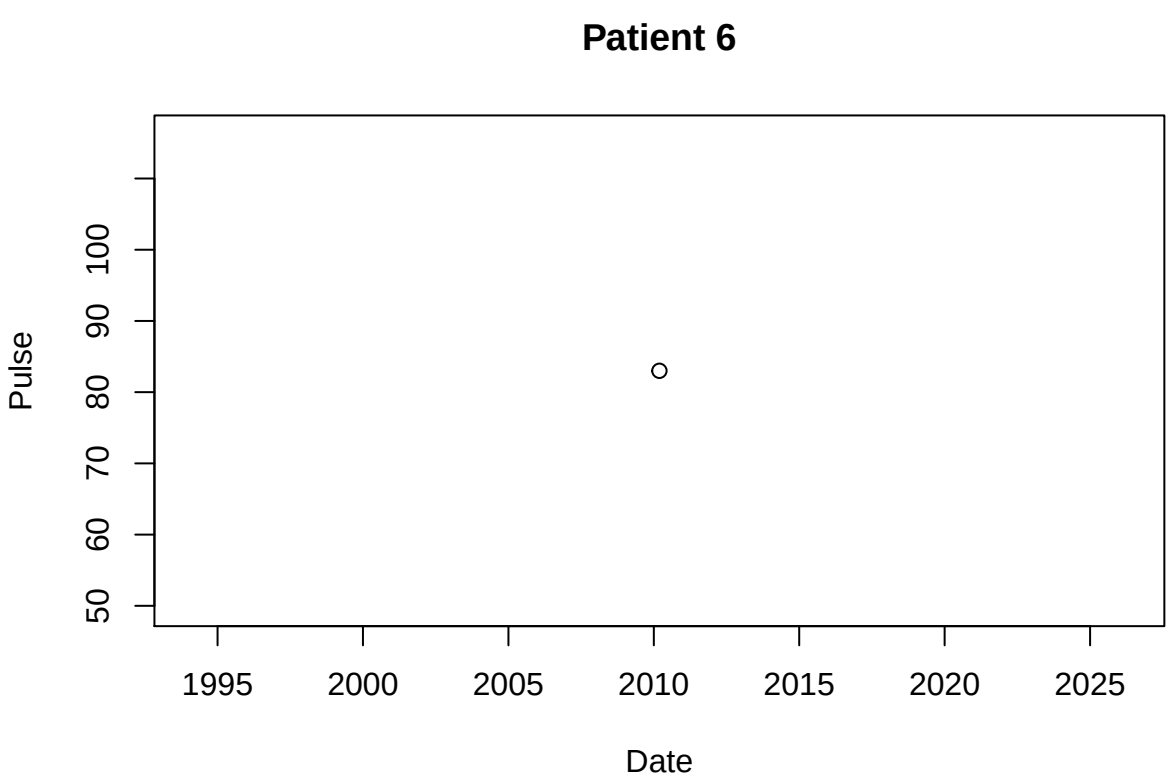

```
##   Date of admission, patient 4 Pulse Temperature Fluid intake
## 1                   2010-04-12    76       98.05         0.65
## 2                   2011-02-16    93       98.26         0.97
## 3                   2012-04-12    96       97.84         0.14
```

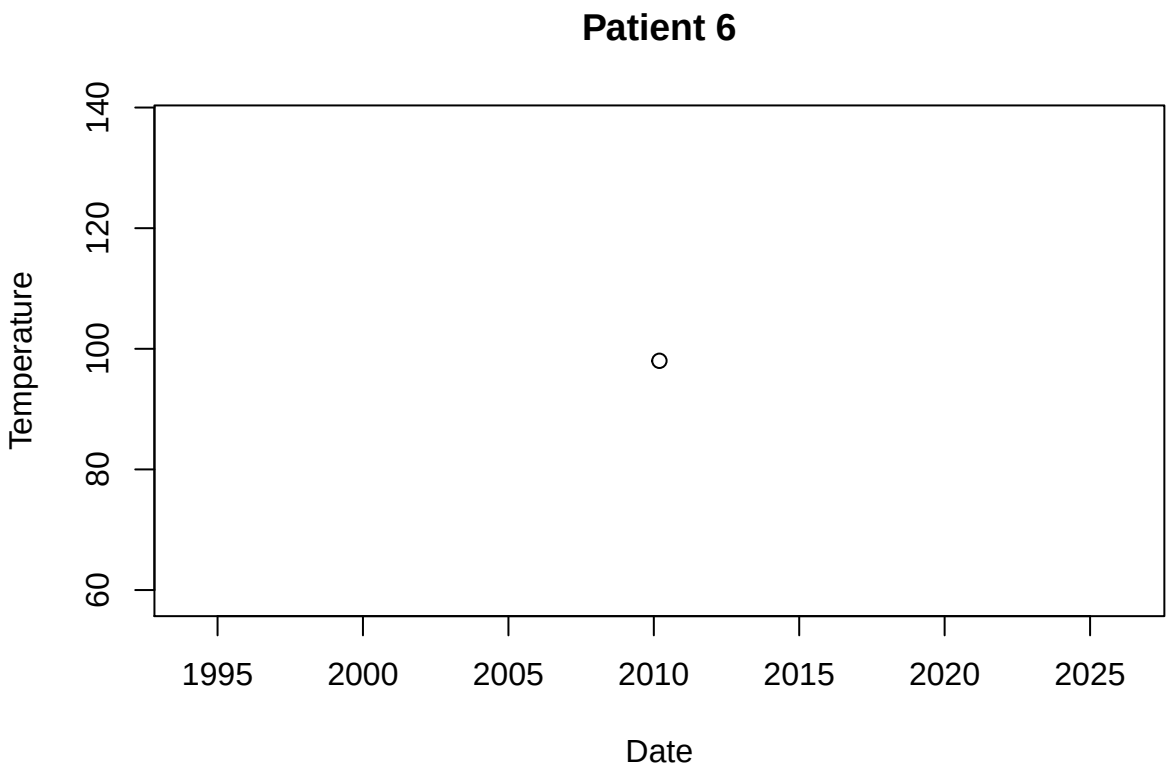

Patient 5
Pulse
90
85
80
75
Mar 15
Apr 01
Apr 15
May 01
May 15
Jun 01
Date

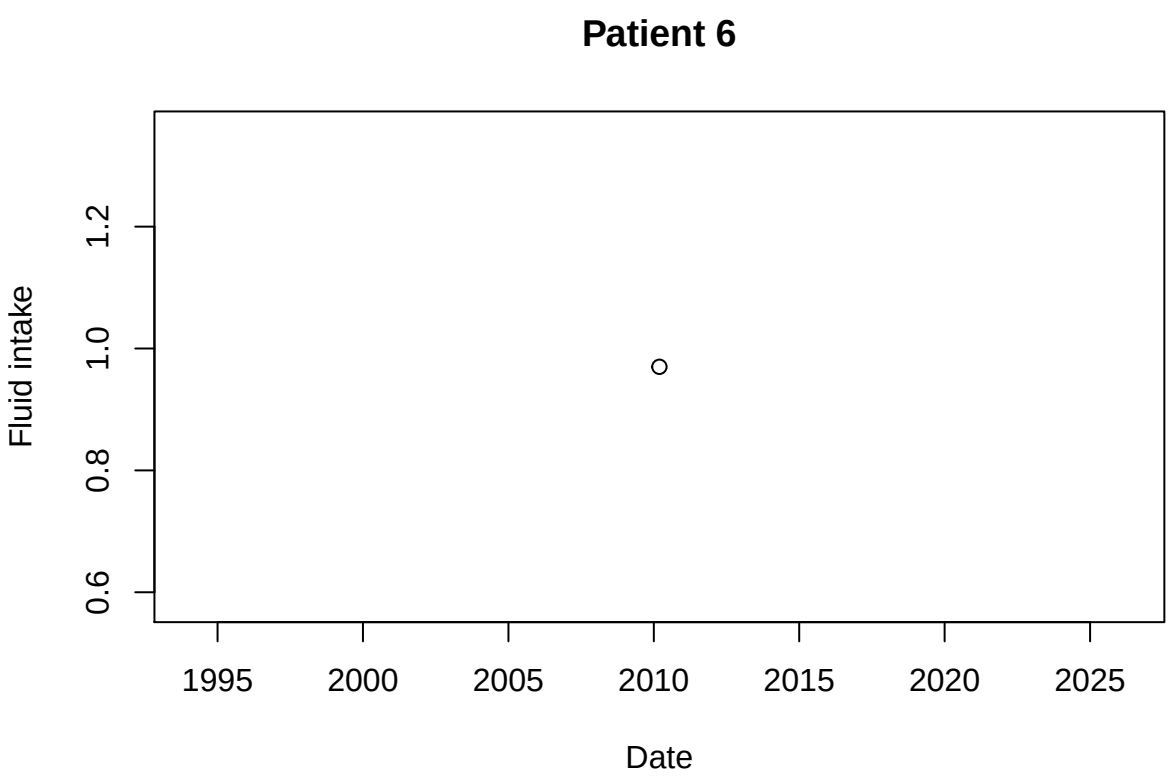

Patient 5
Temperature
99.0
98.5
98.0
97.5
Mar 15
Apr 01
Apr 15
May 01
May 15
Jun 01
Date



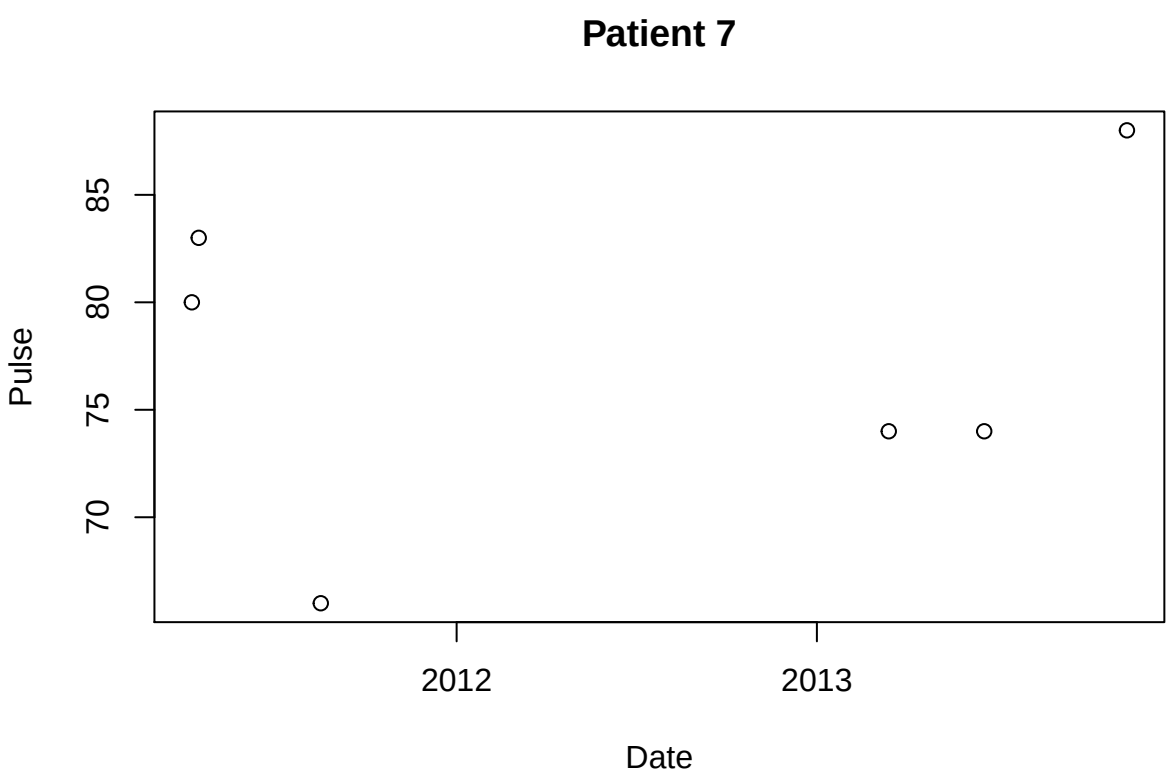

```
##   Date of admission, patient 5 Pulse Temperature Fluid intake
## 1                   2010-03-10    81       99.11         0.66
## 2                   2010-03-25    90       98.58         0.26
## 3                   2010-04-18    75       98.58         0.60
## 4                   2010-06-10    88       97.53         0.29
```

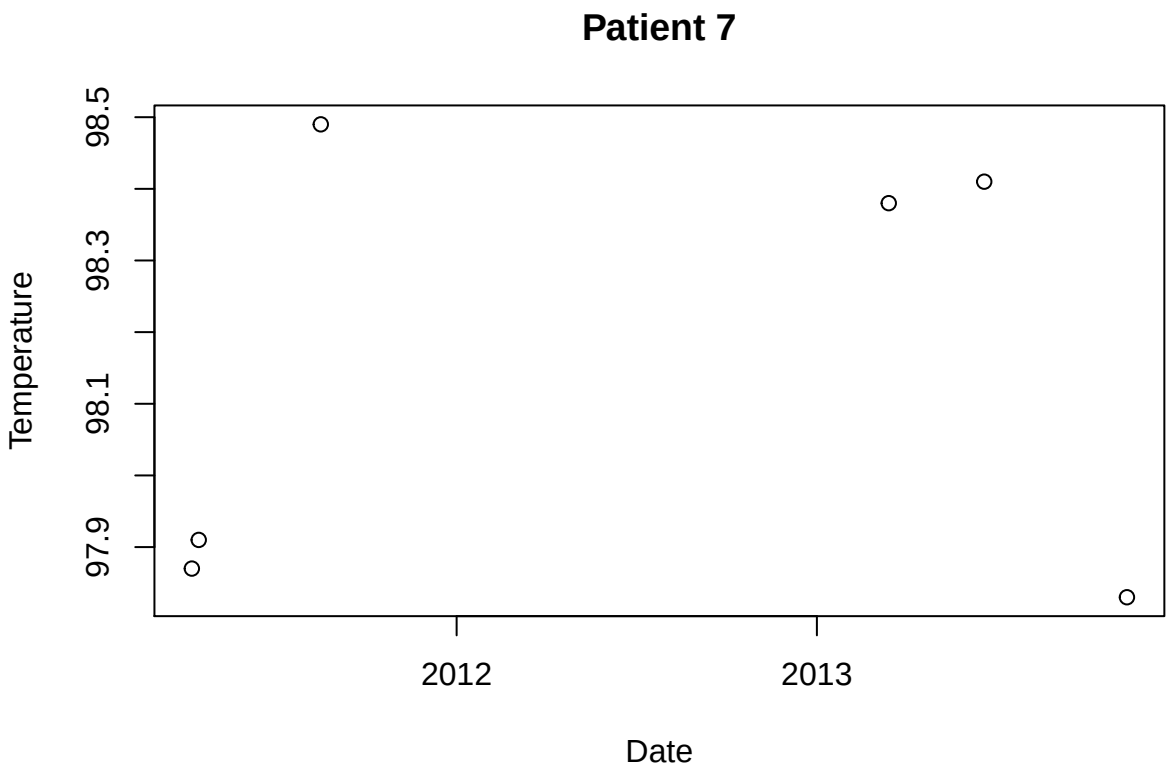

Patient 6
Pulse
50
60
70
80
90
100
1995
2000
2005
2010
2015
2020
2025
Date

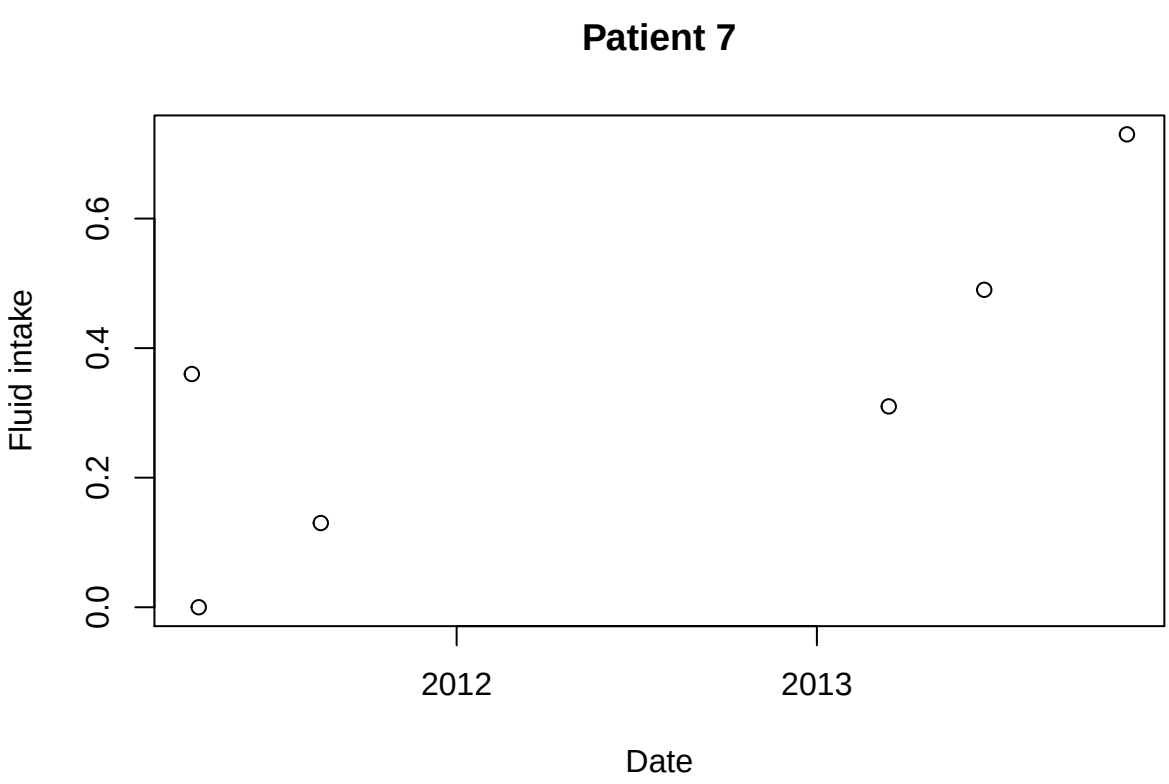

Patient 6
Temperature
60
80
100
120
140
1995
2000
2005
2010
2015
2020
2025
Date



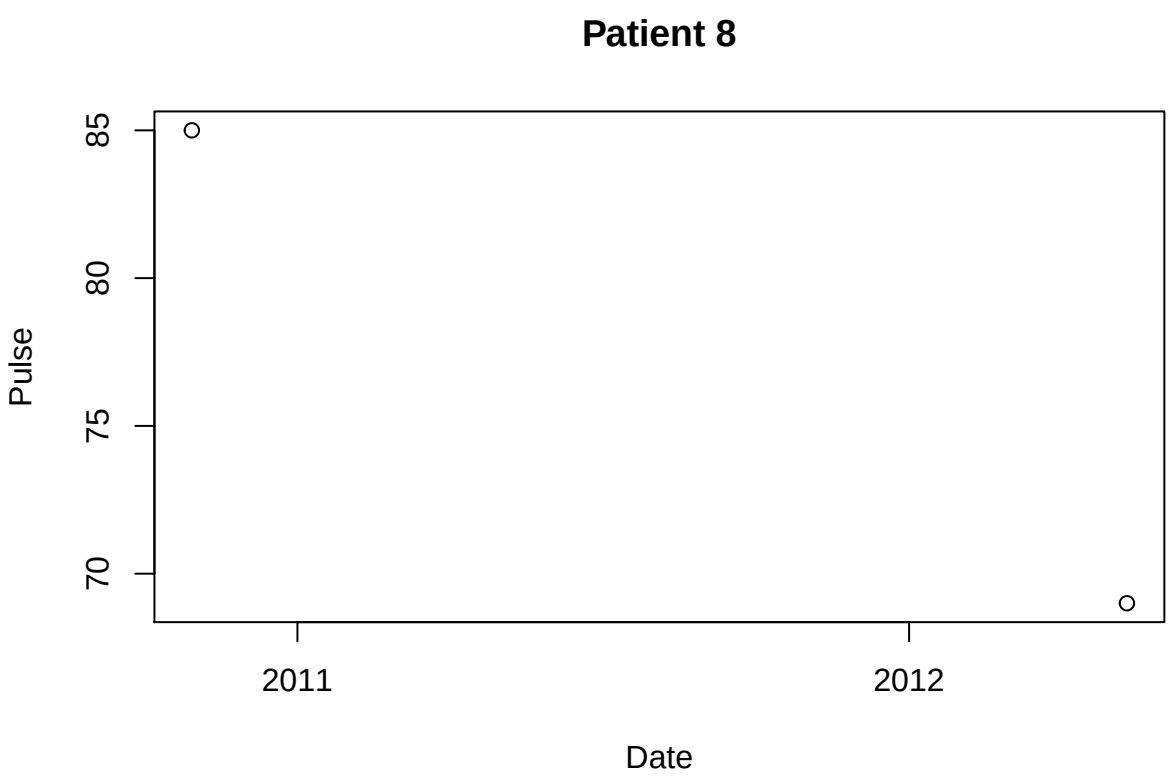

```
##   Date of admission, patient 6 Pulse Temperature Fluid intake
## 1                   2010-03-12    83       98.01         0.97
```

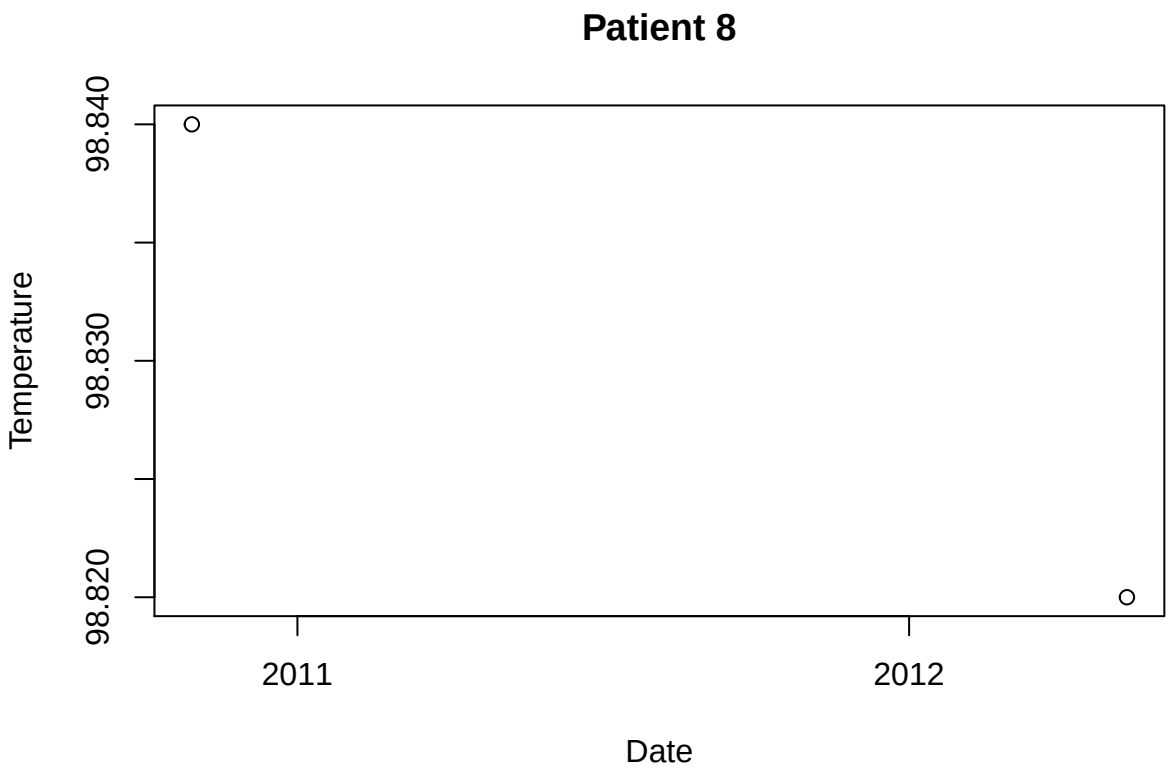

Patient 7
Pulse
70
75
80
85
2012
2013
Date

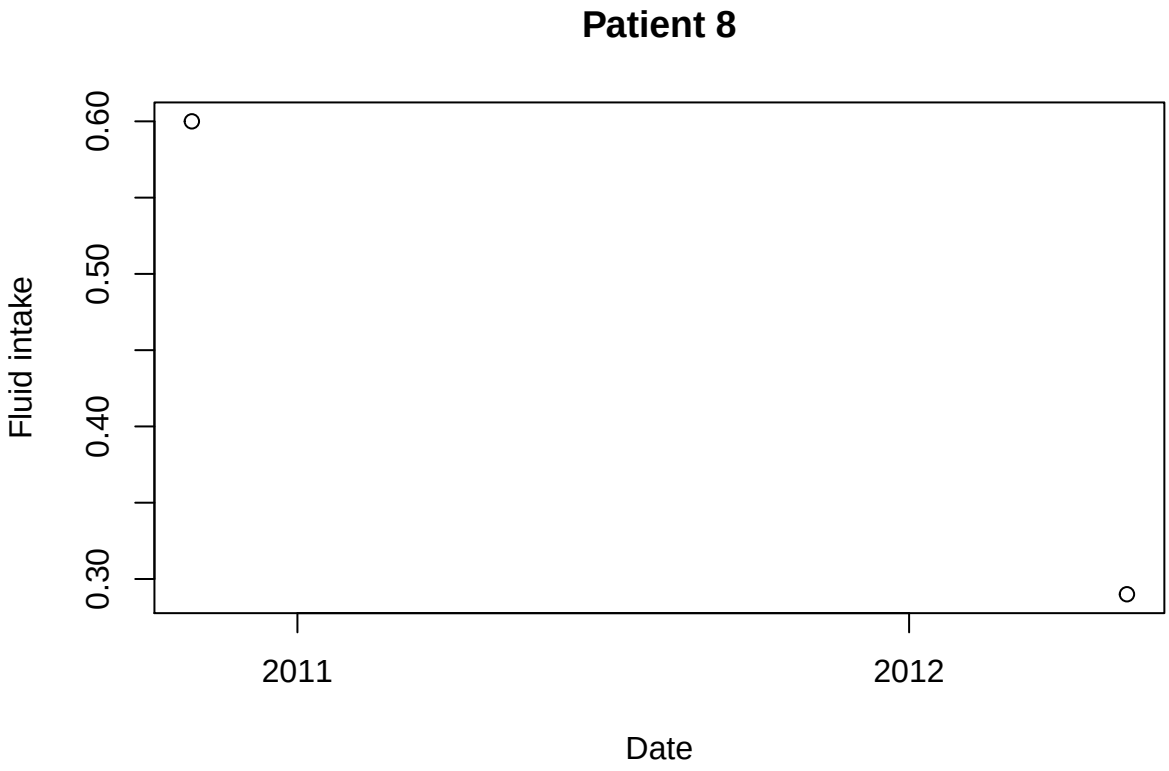

Patient 7
Temperature
98.5
98.3
98.1
97.9
2012
2013
Date



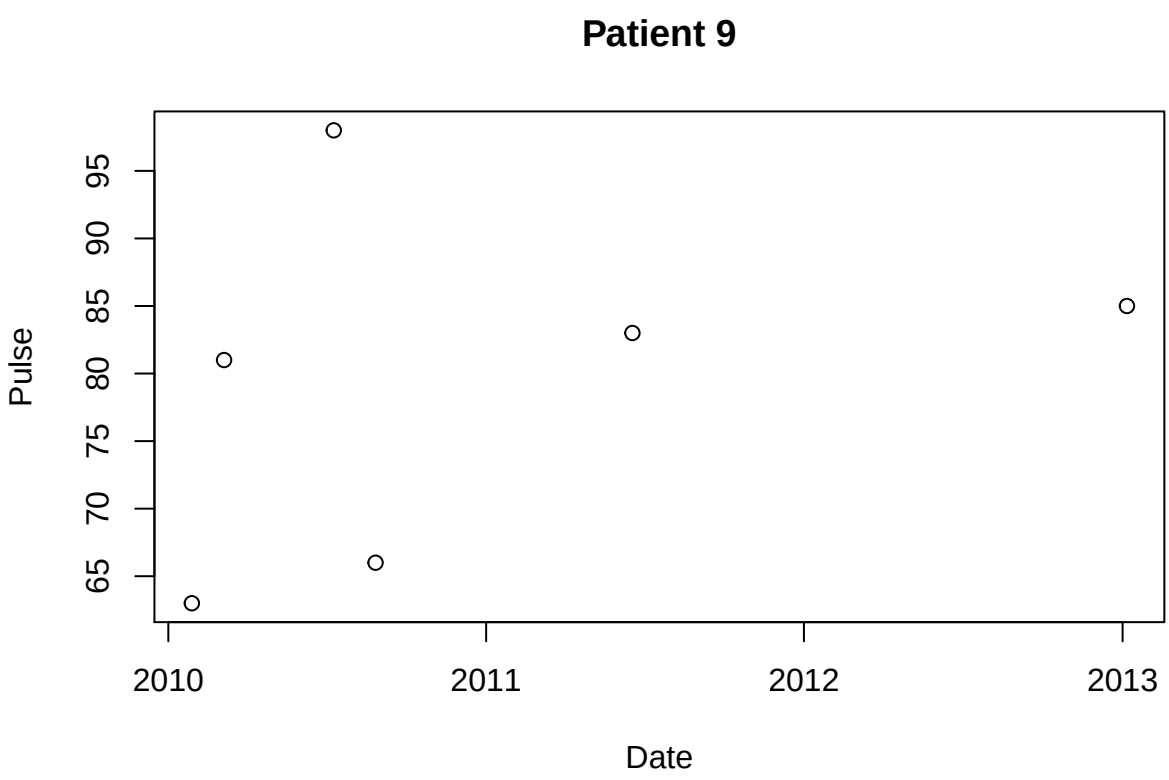

```
##   Date of admission, patient 7 Pulse Temperature Fluid intake
## 1                   2011-04-07    80       97.87         0.36
## 2                   2011-04-14    83       97.91         0.00
## 3                   2011-08-16    66       98.49         0.13
## 4                   2013-03-15    74       98.38         0.31
## 5                   2013-06-20    74       98.41         0.49
## 6                   2013-11-12    88       97.83         0.73
```

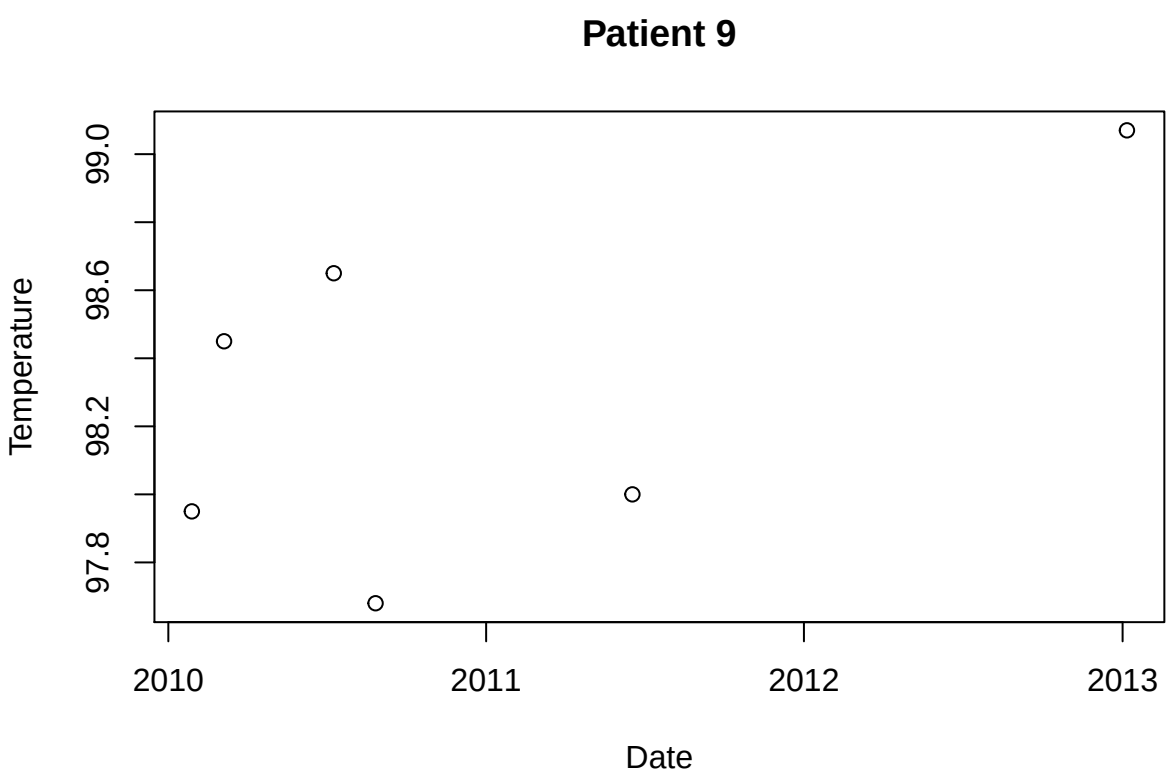

Patient 8
Pulse
85
80
75
70
2011
2012
Date

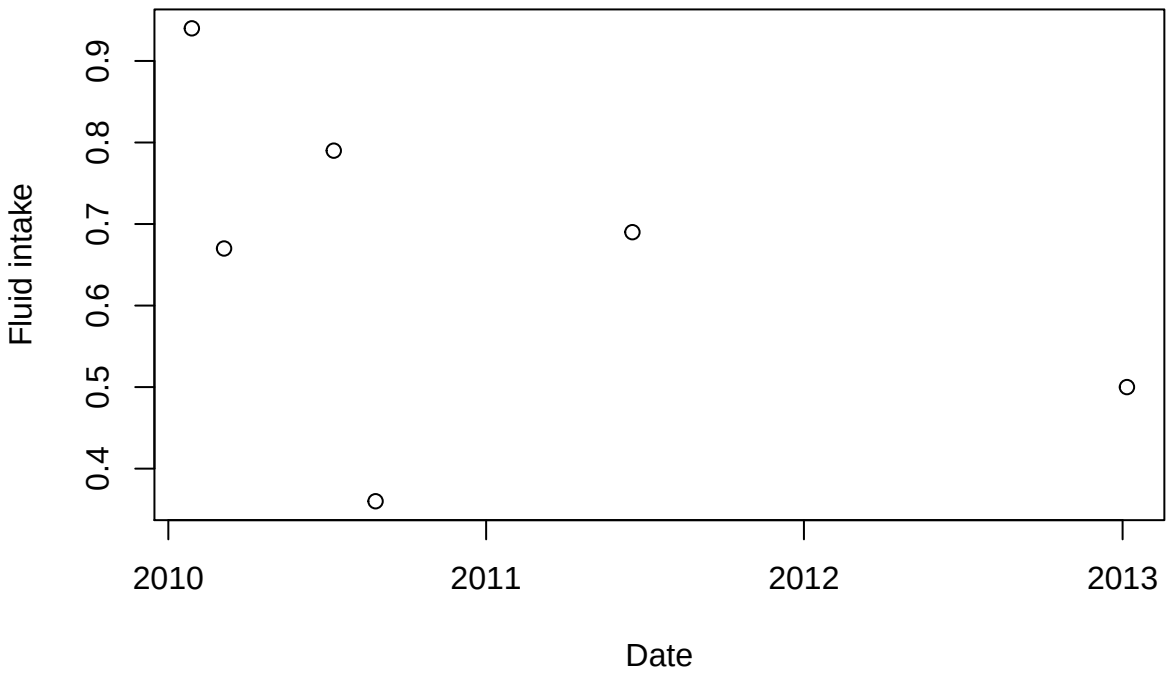

Patient 8
Temperature
98.840
98.830
98.820
2011
2012
Date



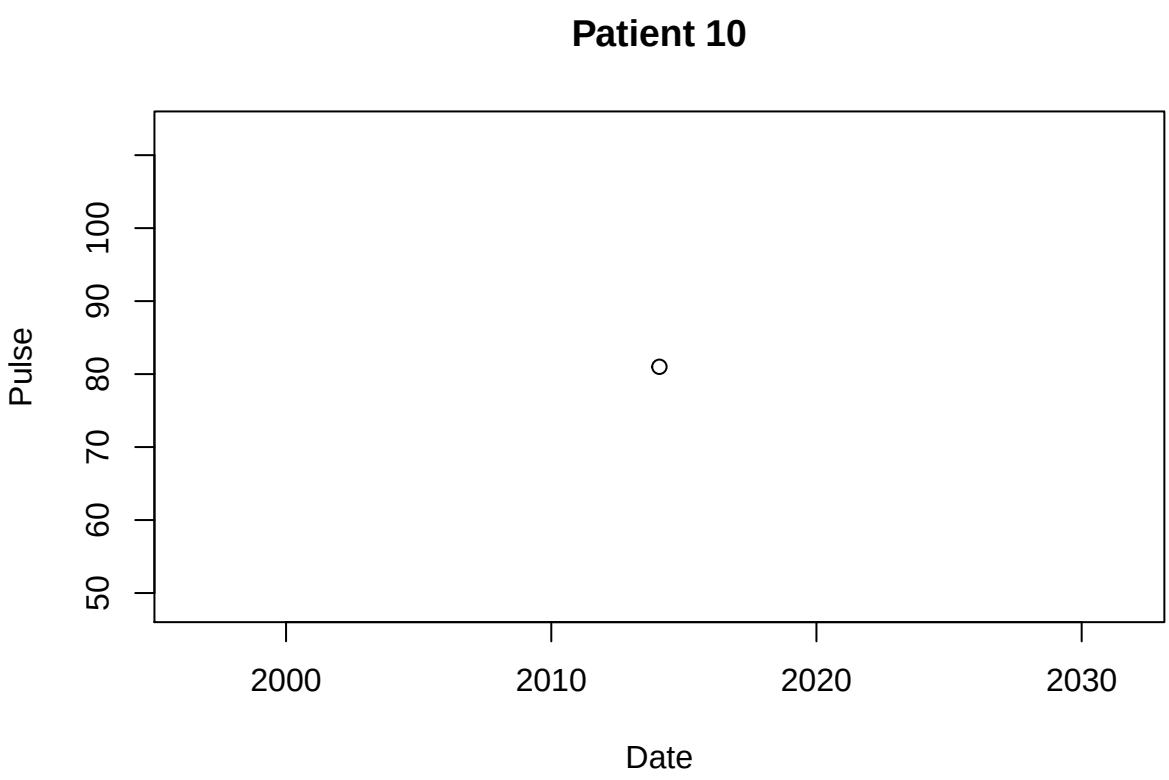

```
##   Date of admission, patient 8 Pulse Temperature Fluid intake
## 1                   2010-10-30    85       98.84         0.60
## 2                   2012-05-10    69       98.82         0.29
```

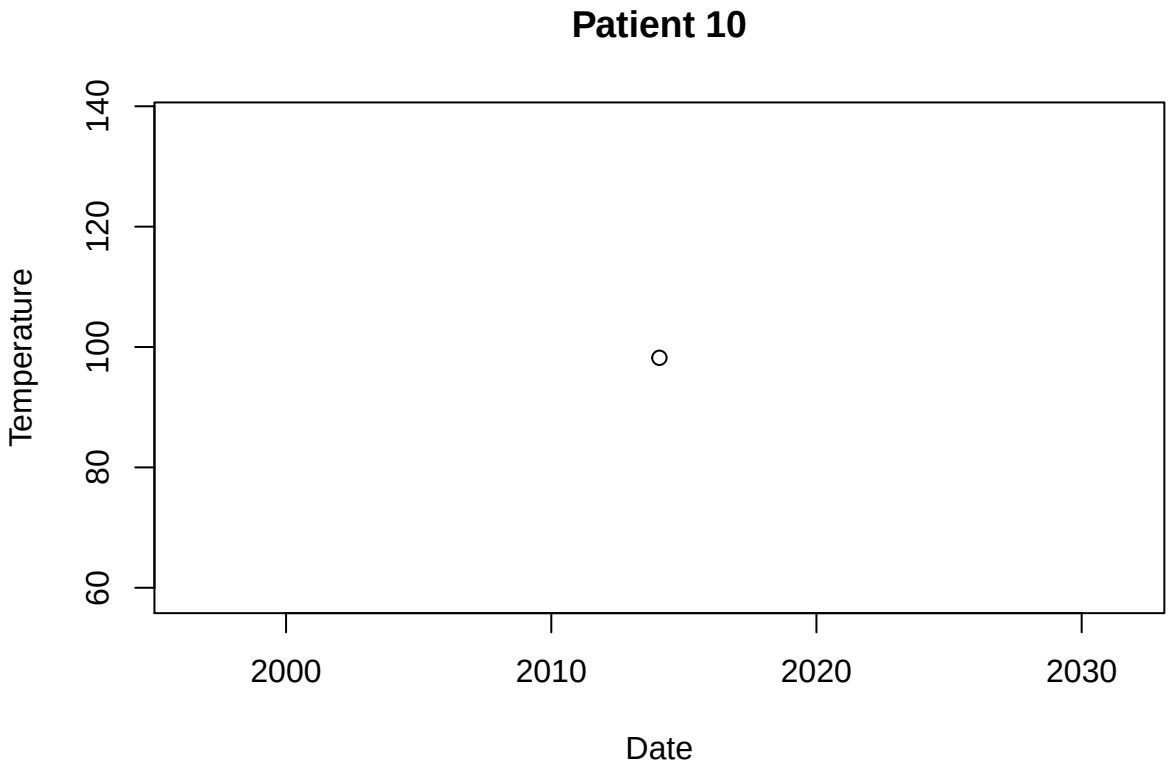

Patient 9
Pulse
65
70
75
80
85
90
95
2010
2011
2012
2013
Date

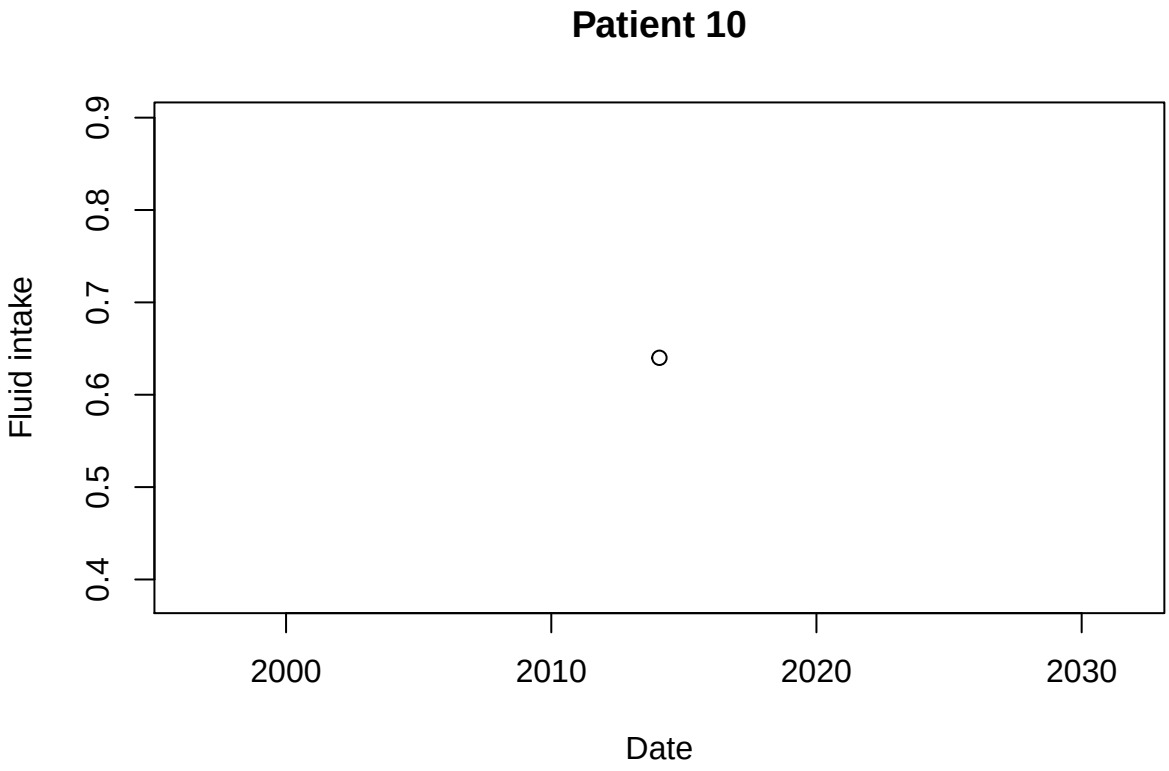

Patient 9
Temperature
97.8
98.2
98.6
99.0
2010
2011
2012
2013
Date




```
##   Date of admission, patient 9 Pulse Temperature Fluid intake
## 1                   2010-01-28    63       97.95         0.94
## 2                   2010-03-06    81       98.45         0.67
## 3                   2010-07-10    98       98.65         0.79
## 4                   2010-08-27    66       97.68         0.36
## 5                   2011-06-18    83       98.00         0.69
## 6                   2013-01-06    85       99.07         0.50
```


Patient 10
Pulse
50
60
70
80
90
100
2000
2010
2020
2030
Date


Patient 10
Temperature
60
80
100
120
140
2000
2010
2020
2030
Date




```
##   Date of admission, patient 10 Pulse Temperature Fluid intake
## 1                    2014-01-29    81       98.21         0.64
```